



PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/204141>

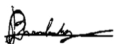
Please be advised that this information was generated on 2020-01-02 and may be subject to change.

The Nurturing Stance

Statement of Originality

This thesis is being submitted to Radboud University and Macquarie University in accordance with the Joint agreement dated 1 September 2017.

To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.



Date: 1 March 2019

The investigations described in this thesis were supported by the Netherlands Organization for Scientific Research (NWO), as part of the research project 'Management of the Self: a Humanities Approach to Self-Management in Psychiatry and Psychosomatic Medicine' Project number: 360-20-360

Copyright
© Daphne Brandenburg, 2019

Cover design by Patrick Lu.
Printed and bound by Book Builders

ISBN:

The Nurturing Stance

Academisch Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken
en
ter verkrijging van de graad van doctor
aan de Macquarie University
op gezag van de Vice-Chancellor professor S.B. Dowton MD

volgens besluit van het college van decanen
in het openbaar te verdedigen op dinsdag 11 juni 2019
om 12.30 uur precies

door

Daphne Desiree Brandenburg

geboren op 24 november 1988
te Heerenveen, Nederland

Promotiecommissie

Promotoren

Prof. dr. J.A.M. Bransen

Prof. dr. M.V.P. Slors

Prof. dr. J. Kennett (Macquarie University, Australië)

Copromotor

Dr. D.W. Strijbos

Manuscriptcommissie:

Prof. dr. C.L.A. Bary (voorzitter)

Prof. dr. M.M.S.K. Sie (Tilburg University)

Prof. dr. G. Meynen (Vrije Universiteit Amsterdam)

Prof. dr. V. McGeer (Princeton University, Verenigde Staten)

Prof. dr. M.S. McKenna (University of Arizona, Verenigde Staten)

Table of Contents

Dutch Abstract	vii
English Abstract	vii
Publications	ix
1.1 Acknowledgements	1
1.2 Chapter Overview	2
2 Addressing the Unaddressed	5
2.1 Three cases	5
2.2 Blaming, excusing, and exempting	7
2.3 Fake it till you make it	10
2.4 The clinical stance	12
2.5 The tripartite theory	15
2.6 Conclusion	17
3 The Nurturing Stance: Making sense of Responsibility without Blame	20
Abstract	29
3.1. Strawsonians and the Clinical Conundrum	22
3.2. Responsibility without Expressed Blame	24
3.3. Responsibility without Blameworthiness	25
3.4. Ambiguous Capacities	27
3.5 The Nurturing Stance	29
3.6 Strawsonians and the Nurturing Stance	33
3.7 Conclusion	36
4 Inadequate Agency and Appropriate Anger	37
Abstract	37
4.1 Communication and cultivation; a response to the blame-sceptic	39
4.2 Attending to reproachful engagements with underdeveloped or compromised agents	41
4.3 An addition to CC accounts	48
4.4 Anticipating an objection	52
4.5 Conclusion	55
5 A Case Study on Strawsonian Exemption from Responsibility	57
Abstract	57
5.1 Strawsonian exemption from responsibility and the coincidence claim	59
5.2 Testing the talk-walk claim	63
5.3 The first challenge: allowing for diversity in moral communication	68
5.4 Testing the walk-talk claim	70
5.5 The second challenge: avoiding undue exclusion from moral conversation	75
5.6 Conclusion	77
6 The Nurturing Dialogue	78
Abstract	78

6.1 Introduction.....	78
6.2 Communicative accounts and recognition of capacity	80
6.3 Communicating with juveniles	83
6.4 Nurturing dialogue.....	85
6.5 Extending recognition of capacity to persons with a mental disability.....	88
6.6 Anticipating an objection	92
Conclusion	95
7 Detached Blame and Nurturing Reproach	96
Abstract.....	96
7.1 Introduction.....	97
7.2 The clinical stance	98
7.3 Nurturing reproach	102
7.4 Comparing nurturing reproach and detached blame	109
7.5 Conclusion	116
8 Conclusion	118
8.1 Summarising conclusion.....	118
8.2 Directions for further research.....	119
References	122

Dutch Abstract

In dit proefschrift bespreek ik hoe we iemand verantwoordelijk kunnen en mogen houden wanneer iemands morele capaciteiten beperkt of onderontwikkeld zijn. Ik richt mij hierbij specifiek op personen die, onder bepaalde omstandigheden, onvoldoende kunnen reageren op de morele redenen die voor een bepaalde handeling spreken, maar wel degelijk kunnen reageren op morele correcties of instructies, en het potentieel hebben om hun gevoeligheid voor de relevante morele redenen te ontwikkelen. Dit kan bijvoorbeeld, onder bepaalde omstandigheden, het geval zijn voor personen met hoog functionerend autisme, mensen met een licht verstandelijke beperking, en kinderen van een bepaalde leeftijd. In dit proefschrift bespreek ik hoe we deze mensen dan verantwoordelijk kunnen houden voor moreel problematisch gedrag.

In het kritische deel van dit proefschrift laat ik zien dat de bestaande literatuur over verantwoordelijkheid deze personen en dit soort situaties over het hoofd heeft gezien. Ik doe dit in twee stappen. Ten eerste laat ik zien dat binnen een hoeveelheid theorieën deze individuen te gemakkelijk geïnterpreteerd kunnen worden – en soms ook expliciet geïnterpreteerd worden – als individuen die geen verantwoordelijkheden kunnen dragen. Ten tweede bekritiseer ik twee theorieën die specifiek betrekking hebben op mensen wiens morele capaciteiten beperkt of onderontwikkeld zijn, respectievelijk: Hanna Pickard's *'Clinical Stance'* en David Shoemaker's *'Tripartite Theory'*. Ik laat zien dat ook deze theorieën niet volledig uit kunnen leggen hoe de personen waar ik mij op richt verantwoordelijk kunnen worden gehouden voor moreel laakbaar gedrag.

In het positieve deel van het proefschrift ontwikkel ik een normatieve theorie betreffende onze omgang met deze individuen. Ik noem deze theorie de *'Nurturing Stance'*. Vanuit een *Nurturing Stance*, achten we iemand niet laakbaar voor zijn of haar problematische gedrag, omdat we de persoon nog niet voldoende instaat achten om de relevante morele redenen te registreren onder bepaalde omstandigheden. Maar vanuit deze houding zien we de persoon wel als iemand met wie we een moreel gesprek aan kunnen gaan op basis van dit problematische gedrag, en ook als iemand die naar aanleiding van dit gesprek bepaalde verantwoordelijkheden kan nemen. We vragen de persoon om te participeren in een proces dat hem of haar in staat stelt morele redenen in de toekomst te registreren en/of na te leven. Het is dan iemands verantwoordelijkheid om – voor zover diegene daartoe in staat is – stappen te nemen die helpen vergelijkbaar problematisch gedrag in de toekomst te voorkomen. Ter afsluiting bespreek ik hoe mijn theorie in twee institutionele contexten van belang kan zijn, ten eerste in het strafrecht en ten tweede binnen de psychiatrische kliniek.

English Abstract

This thesis provides an account of how we may hold people responsible when their capacities to abide by moral norms are underdeveloped or compromised. It focuses on individuals who cannot yet sufficiently respond to moral reasons in particular types of circumstances, but who 1) can become able to do so, and who 2) are receptive to moral instruction and moral correction. For example highly functioning autistic individuals, people with mild intellectual disability, and children of a certain age will –under certain conditions and in particular types of circumstances- meet these criteria.

The negative part of this thesis exists in two parts. Firstly, I will illustrate how this group of individuals has been commonly overlooked or misrepresented in the responsibility literature. I show how on a number of accounts they may be mistakenly interpreted -or are even mistakenly interpreted- as individuals who are insensitive to moral reasoning and entirely exempted from responsibility. Secondly, I will critically engage with two recent accounts of the responsibility of individuals whose agency is diminished: Hanna Pickard's clinical stance, and David Shoemaker's tripartite theory of responsibility. Hanna Pickard argues that we should consider people with personality disorder to be responsible and blameworthy, but should refrain from emotionally blaming them in therapeutic and legal contexts. David Shoemaker argues that individuals with high functioning autism are unable to respond to moral reasons and that our responses to them should therefore target non-moral abilities and character traits. I will take issue with both accounts, and demonstrate why further theorizing is needed.

The positive part of this thesis exists in the development of my own account called the nurturing stance. From a nurturing stance, one does not attribute blameworthiness to a person or blame a person for her objectionable conduct. From this stance perceived objectionable conduct is understood to give rise to future responsibilities that we may hold the person to upon witnessing her failure to meet an interpersonal expectation. We urge the person to take (or keep taking) responsibility for developing and repairing their own compromised or underdeveloped abilities. This is to say that we expect them to do what they can to participate in, and contribute to a process of development or recovery. When one is subject to the nurturing stance, one is reasoned with and called upon. So by adopting a nurturing stance, one importantly recognizes and respects the other person as someone who one can reason with and who can come to meet the shared expectations and norms at stake. I pay special attention to possible applications of this stance in 1) criminal law and 2) psychiatric practice.

Publications

Most of the chapters included in this thesis are reproductions of published and unpublished papers, as specified below:

Chapter 3: Brandenburg, D. (2017). The Nurturing Stance: Making Sense of Responsibility without Blame. *Pacific Philosophical Quarterly*, 99(S1). <https://doi.org/10.1111/papq.12210>

Chapter 4: Brandenburg, D. (2019). Inadequate Agency and Appropriate Anger. *Ethical Theory Moral Practice* <https://doi.org/10.1007/s10677-019-09982-w>

Chapter 5: Brandenburg, D. (submitted) A Case Study of Strawsonian Exemption from Responsibility.

Chapter 6: Brandenburg, D. & Strijbos, D. (under review) Detached Blame and Nurturing reproach.

Chapter 7: Brandenburg, D. (submitted) The Nurturing Dialogue.

Other peer-reviewed publications

Brandenburg, D. (2016). Implicit attitudes and the social capacity for free will. *Philosophical Psychology*, 29(8). <https://doi.org/10.1080/09515089.2016.1235263>

Brandenburg, D. (2018). De bemoedigende houding. *Algemeen Nederlands Tijdschrift voor Wijsbegeerte*. 110 (4). <https://doi.org/10.5117/ANTW2018.4.002.BRAN>

1.1 Acknowledgements

My first thanks go to the psychiatrists, psychologists, and nurses who selflessly made the time to let me interview them. The anecdotes and insights I have been offered by these experts from a third line clinic for autism at *Dimence*, have greatly increased my familiarity with the practices that are analysed in this thesis. Special thanks go to Derek Strijbos and Gerrit Glas for, in addition to sharing their experiences with me, enabling me to do fieldwork at psychiatric wards and for putting me in touch with other clinicians.

I would not have been able to write this thesis without the guidance of my supervisors. Jan, during our brainstorm sessions you always enabled me to identify the issues that truly interest and move me. Thank you for setting me on the path that led to this body of work. Marc, your unrelenting enthusiasm and criticism always motivated me to always go that extra mile. Thanks also for having read every single version of every single paper (there have been so many!). Derek, thank you for your invaluable input from a psychiatric angle. Jeanette, you were already an inspiration to me before becoming my supervisor. It has been an honour to work with you and I hope I may become a philosopher like you one day.

Berrie, Els, Marcella, and Renee, we have been friends for so long I can't even imagine how to achieve anything without being able to talk to you whenever from wherever. Patrick, Alisha, Anissa, and Yves, I am so glad to have met you in Sydney. You have always been there to discuss my thesis and, more so, to distract me from it by the end of the week. Patrick, thank you for designing the awesome cover page, you are an amazing artist. Max and Wei, thanks for being my PhD coffee team throughout these last two months. Thanks, also, to coffee. I could not have done this without you. Nor could I have done this without all the friends that I have forgotten to give credit to here.

My sister and her husband deserve special mention simply because they are the best, but also because they make sure I file my tax-returns, remind me of family gatherings, and always have an endless supply of snacks in their house. Last but not least, I want to thank my parents Annemieke and Chris, who made it all possible by having me, and always supporting me in so many different ways. I dedicate this thesis to them. They are also causally responsible for any of the mistakes I have made.

1.2 Chapter Overview

This thesis is motivated by the observation that philosophical literature on responsibility has neglected a relevant sphere of our human interactions. The second chapter will begin with bolstering the claim that a satisfactory account of these types of valuable human interactions is, as of yet, not available. I illustrate a valuable manner of responding to individuals who behave objectionably, but who are not considered blameworthy because they are not (yet) independently able to understand and respond to moral reasons to act differently under similar circumstances. After that I will discuss why these responses are different from any response to objectionable conduct that has so far been conceptualised in the literature on moral responsibility.

The third chapter conceptualises the under-theorized stance we take towards those whose abilities are underdeveloped or compromised. It responds to Hannah Pickard's observation that the Strawsonian literature would have trouble accounting for the fact that clinicians report that they hold patients responsible for morally objectionable behaviour but at the same time consider blaming attitudes to be inappropriate. I discuss how both Pickard and Strawsonians fail to fully explain the psychiatric practices at stake and conceptualise, what I call, a *nurturing stance* towards underdeveloped or compromised individuals that can explain them. From this stance we do not attribute blame or blameworthiness to a person for objectionable conduct because we do not consider this person to (yet) be independently capable to act differently in similar circumstances. But we do consider the person to be someone who we can morally reason with about this objectionable conduct, and to be someone who can take future directed responsibilities because of this objectionable conduct.

In the fourth chapter I further elaborate on the nurturing stance and discuss the theoretical implications it has for functional accounts of blame. Communication and cultivation accounts of responsibility (CC accounts) argue that blaming has an important communicative and agency-cultivating function when addressed at someone we consider to be deserving of blame. On these accounts, responsible agents are agents who can understand blaming attitudes and are (generally) sensitive to their moral-agency cultivating function. I examine our reproachful engagements with agents whose moral agency is underdeveloped or compromised. I discuss how these engagements compare to blaming on CC accounts and argue that reproachful engagements can have an important communicative and agency-cultivating point. The nurturing stance provides an addition to CC-accounts that explains how underdeveloped and compromised

agents can be held responsible. I will show how this addition resolves ambiguities in accounts from McKenna, Vargas and McGeer. I conclude the chapter by anticipating an objection CC accounts could raise: that reproach and blame are co-extensive.

In the fifth chapter I discuss theoretical implications for Strawsonian accounts of exemption from responsibility. I focus on David Shoemaker's account of exemption from accountability. The nurturing stance highlights that individuals who are not considered blameworthy due to compromised agency may nevertheless be sensitive to moral address. Strawsonian exemption is generally characterized by the suggestion that insensitivity to moral address robustly tracks or constitutes impaired reason-responsiveness. In this chapter I ask what the autistic deficit in affective attunement implies for both a person's sensitivity to reactive attitudes as moral address, and a person's responsiveness to moral reasons, and thereby test the purported connection between these two capacities. I argue that, in order to avoid unduly excluding individuals with high functioning autism from the moral community, Strawsonians need to adhere to a liberal notion of sensitivity to reactive attitudes, and a liberal notion of reason-responsiveness. But doing so undermines the theoretical import of a constitutive or tracking connection between these capacities. In line with the last chapter, I suggest the connection be reconsidered as a developmental connection.

The sixth chapter discusses how the nurturing stance may help extend a call for recognition within the context of criminal law. Article 12 of the United Nations Convention on the Rights of Persons with Disabilities calls for recognition of the legal capacity of all people with mental disability in the context of civil law. When this call for recognition is extended to criminal law it seems to imply we should reject incapacity based defences for persons with mental disabilities. But if it is, due to a person's mental disability, difficult to impossible for this person to abide by the law it seems unfair not to take this into account. This raises the question whether states can respond to a call for wider recognition of capacity in the context of criminal law without unfairly blaming and punishing persons with mental disability. In this chapter I argue that a modified version of the communicative theory of criminal law can provide such an answer.

The seventh chapter discusses the implications of the Nurturing Stance for psychiatric practice. It discusses a qualitative study with mental health care clinicians that preceded and informed the development of my account. The outcomes of this study are compared with Hannah Pickard's account of responsibility without blame. Hannah Pickard has provided an account of how service users should be held responsible for objectionable conduct within contexts of mental health care. Her account called 'the clinical stance' is based on what are

considered best practices in the treatment of people with Borderline Personality Disorder. We compare this stance to what are considered best practices at a specialized clinic for people with Autism Spectrum Disorder in the Netherlands. We refer to our alternative account of best practices as ‘nurturing reproach’ and discuss relevant similarities and differences between the clinical stance and nurturing reproach.

In conclusion I provide a brief summary and identify directions for further research.

2 Addressing the Unaddressed

A literature review

This thesis is motivated by the observation that philosophical literature on responsibility has neglected a relevant sphere of our human interactions. I argue that it has left unaddressed or sometimes even misrepresented a valuable way in which we can respond to the objectionable conduct of a person whom we deem insufficiently competent to be blameworthy. By conceptualising and evaluating this form of relating, I hope to provide an account that allows for greater recognition and inclusion of compromised or underdeveloped agents as members of a moral community, and facilitates moral development. To keep things in a logical order, this chapter will first bolster the claim that a satisfactory account of these types of human interactions is not yet available. In the next section I will provide three cases that are illustrative of the type of response to others that I have in mind. After that I will discuss why these responses are different from any response to objectionable conduct that has so far been conceptualised in the literature on moral responsibility.

2.1 Three cases

Consider what happens in the following scene from a Louis Theroux documentary about people with severe brain damage. Among the persons with brain damage who are followed in this documentary is a man called Earl. Earl had been in a car accident, in which his friend died. He sustained serious brain damage and the court deemed him unfit to plead. At the time of filming, Earl is living in a neurobehavioural rehabilitation centre in Leeds, unable to live independently. It is said that before the accident he was a charming young man, but currently he is sometimes hostile towards others, especially towards his mother. During one scene in the documentary, Louis is talking to Earl's mum when Earl suddenly comes bursting in and is verbally abusive to her. His mum starts crying and tells Louis that this abuse happens all the time. Louis looks at Earl and asks him if he has just been playing FIFA and lost. Earl confirms that he has. Louis then asks, 'You are taking this out on your mum, aren't you?' Earl doesn't respond. Louis adds: 'I think you should apologise to your mum.' Earl asks, 'Why? She doesn't care.' Louis then looks demonstratively at his mum (who is crying). Earl also looks at his mum, and then walks over, apologises, and gives her a big hug (Pickup & Theroux 2016).

Something relevantly similar happens in the following scenario, which is loosely based on a court case that was heard in Melbourne (Parish v DPP, 2007). A woman called Lillian took a

train into the city to meet up with her sister. On the train, a man called Jake sat down in front of her. He put his knee against hers. She moved away. He then moved closer to her and started brushing her knee. Lillian, feeling increasingly uncomfortable, turned away from him, ignored the brushing, and stared out of the window. Later, when Lillian got off the train, Jake followed her up some stairs and rubbed her lower back and the top of her buttocks. When Lillian arrived at her sister's place she broke down in tears and told her what had happened. Together they decided to report Jake to the police.

It turned out that Jake has Asperger's. After listening to Jake and to the expert reports, the court ruled that Jake was not guilty because, due to his autism, he was not aware and did not know that Lillian did not consent to his actions. Lillian and Jake got to speak to each other through a restorative justice programme.¹ Lillian told Jake how his behaviour had upset her and asked him what he planned to do now that he *did* know how his behaviour had affected her and may affect other women if he were to do something similar in the future.

Consider also, lastly, the dialogue between two people I'll call Rob and Anika that takes place after Anika had gone through a period of mania. When she was manic, Anika went to Rob's house in a skimpy outfit and made explicit sexual advances towards him in front of his partner and children. She also repeatedly insisted on holding the newborn baby but was reckless when she held her, and would have dropped her if Rob hadn't intervened.

After Anika's manic period, Rob and Anika met up for lunch. Rob told her what had happened and how worried he and his partner had been when she almost dropped the baby. Anika was unaware of how explicit and risky her conduct had really been. After Rob told her what had happened she was very distressed and apologised profusely. Rob comforted her and said he understood that she did not mean to be careless and that she had not been in control at the time. He did not hold it against her. He stressed that he was telling her this because he wanted her to know that her bipolar symptoms were worsening and that this constituted a possible danger to her and to others. He asked what she thought could and should be done in order to try and stop similar situations from occurring.

I believe these three cases are illustrative of a type of stance people may take towards an individual who has behaved objectionably but whom one does not consider blameworthy due to incapacity. I will refer to this stance as the *nurturing stance*. Two features characterise the response that is expressive of this stance. First, this response involves engagement of the other person in practical reasoning about the objectionable conduct that took place or is taking place. Second,

¹ This part of the scenario is fictional; in the actual court case no restorative justice programme was available.

this response involves an appeal to the person to be involved in a process of change that can help stop or diminish the objectionable conduct now or in the future.

The problem, as I see it, is that the categories that are currently on offer in the responsibility literature do not describe or evaluate this response. A related problem is that the current literature thereby sometimes suggests that we are to respond to these individuals in ways that are problematic. If I am correct, the moral landscape needs to be carved out differently in order to account for these responses. The next section discusses why these responses to objectionable conduct do not fit into the traditional categories of blaming, exempting, and excusing. Section 3 then rejects the option that these responses should be understood as ‘fake it till you make it’ responses. Sections 4 and 5 discuss two recent accounts that explicitly address compromised agency and responsibility: Hanna Pickard’s clinical stance and David Shoemaker’s tripartite theory. I explain why these respective accounts also do not address the specific responses that are of interest here. I conclude with a diagram that illustrates where this thesis will contribute to the field.

2.2 Blaming, excusing, and exempting

The most commonly discussed responses to morally objectionable conduct are blaming, excusing, or exempting. Do the responses from the three scenarios qualify as any of those three responses?

What does it mean to blame a person? And can we say that the protagonists in our examples are blaming? Philosophers widely disagree about what blame exactly amounts to (Coates & Tognazzini, 2012). Some define blame in terms of the judgements we make when we blame someone (Hieronymi, 2004). For others blame exists in a particular type of negative emotional response to a person. Among others, the emotions that have been said to constitute blame are resentment, indignation, moral anger, or hostile feelings (Menges, 2017; H. Pickard, 2013; Shoemaker, 2015; e.g. Peter F. Strawson, 1962; Wallace, 1996). Other accounts propose that blame is a particular belief–desire pair (Scanlon, 2010; Sher, 2008), while alternative accounts suggest that blame is defined by its function (McGeer, 2014; e.g. M. McKenna, 2012; Vargas, 2013a).

Although there are many different definitions of blame, the responses that are central to the three cases above would not count as instances of blame on the great majority of them. On most accounts, blaming minimally involves two (possibly erroneous) attributions to the blamee, which are absent in the cases discussed. First, what we say of the person when we blame a person is that her action expressed bad quality of will. That is, we perceive the person to be

someone who displayed insufficient regard for the interests and concerns of others. Second, when we blame someone, we also assume that this person has sufficient capacity to independently respond to the moral reasons that she disregarded. It is generally suggested that one can only truly have a bad quality of will if one is already responsive to the moral reasons that were disregarded.

But when Louis, Lillian, and Rob respond to Earl, Jake, and Anika respectively, they do not attribute sufficient responsiveness and bad quality of will to them. To the extent that blaming involves attribution of bad quality of will and sufficient responsiveness to moral reasons, the responses we are interested in here therefore do not qualify as blame.

On some accounts of blame we can attribute bad quality of will to a person without also attributing sufficient capacity. Certain attributionist accounts of blame amount to a negative evaluation of another person's traits or character where this does not require that the person also has capacity to respond to the moral reasons that were disregarded. One can, for example, consider a psychopath to have morally problematic attitudes and commitments without attributing sufficient reason-responsiveness to the psychopath (Talbert, 2008). It is contested whether such responses should be conceived of as forms of holding responsible. But what matters here is that the responses in our cases do not reduce to such attribution of bad character traits either.

Neither would these responses qualify as blaming on most emotional accounts of blame. The ways in which the protagonists in these examples are affected by the other person's conduct differ from the range of emotions that philosophers have argued can constitute blame. To the extent that an appraisal of bad quality of will is involved in the emotional response that is defined as blame, these responses would not qualify because such appraisal is absent. More generally, it would simply be inaccurate to describe the protagonists from our examples as resentful, indignant, or hostile towards the other person (see chapter 4 for an elaboration of this point).

In chapter 4 I pay special attention to functional accounts of blame. Some revisionist functional accounts of blame could perhaps conceive of the responses in my examples *as* forms of blaming, because these responses fulfil the function that they identify with blame (e.g. McGeer, 2018). I discuss why it would be better to maintain that the responses above are not blame, and explain how they could instead be incorporated into a functional account of responsibility.

If the cases discussed here are not instances of blaming, how else may we describe this way of relating to individuals who behave objectionably? Traditionally, if a person's skills are underdeveloped or compromised in such a way that she is not blameworthy for her

objectionable conduct, the person is either considered exempted or excused. However, I contend that the responses that are central to this thesis are poorly described as *practices* of exempting or excusing.

In 'Freedom and resentment', P. F. Strawson writes that there is a particular class of pleas that indicate a person is 'abnormal or morally underdeveloped' (P. F. Strawson, 2008, p. 9). These pleas invite us to take an objective stance towards the person. Strawson writes that the objective stance

cannot include the range of reactive feelings and attitudes which belong to involvement or participation with others in inter-personal human relationships; it cannot include resentment, gratitude, forgiveness, anger, or the sort of love which two adults can sometimes be said to feel reciprocally, for each other. If your attitude towards someone is wholly objective, then though you may fight him, you cannot quarrel with him, and though you may talk to him, even negotiate with him, you cannot reason with him. You can at most pretend to quarrel, or to reason, with him (P. F. Strawson, 2008, p. 10).

Within the Strawsonian tradition, the practice of exempting a person is associated with taking such an objective stance towards a person. The protagonists in our examples are not exempting the other person in this sense because they do not take such an objective stance towards the person. They *are* reasoning with the other person, and, in some way, are involved and participating with this other person in an interpersonal human relationship (see chapter 3 for an elaboration of this point).

Those who are not Strawsonians could maintain that the persons from our examples are exempted from blame for their failure to abide by moral rules on the basis of their incapacity, without also implying that exempting a person amounts to taking an objective stance towards this person (Kennett, 2007). However, even then it is hard to see how the responses we are concerned with would be a *practice* of exempting. Louis, Lillian, and Rob do not *only* exempt a person from responsibility for the harm done on this occasion. They also make an appeal to the other person.

Could we instead say that they are excusing the other person? A number of philosophers would say that a person is excused rather than exempted when their incapacity is local or temporary (e.g. Brink & Nelkin, 2013). Furthermore, on a social model of disability, the distinction between a circumstantial defence (an excuse) and an incapacity-based defence (an exemption) is unclear. Addressing this issue is beyond the scope of this chapter, but I do say a bit more about it in chapter 5.

Let us accept, for the sake of argument, that Anika and Jake from our examples should be excused from blame (rather than exempted from responsibility) on the basis of their

respective temporary and local incapacity to register moral reasons. Especially in the case of Anika, one could plausibly say that Rob *is* excusing her when he tells her he does not hold it against her that she acted as she did. But again, this is not all that he does. Anika is not completely off the hook; he appeals to her to do something because of what has happened.

Recall that the aim was to make sense of the fact that these protagonists engage the other person in practical reasoning and appeal to them to constructively respond to what has happened. This particular response cannot be adequately described as a practice of blaming the person, exempting the person, or excusing the person.

How should we conceptualise and evaluate this particular response, then? As of yet, the answers are unclear. Are there other categories available that do conceptualise and evaluate this type of response to objectionable conduct?

2.3 Fake it till you make it

Philosophers sometimes mention in passing that we typically relate to children who act objectionably ‘as if’ we hold the child responsible. The thought is that we respond like this because the child is not yet sufficiently capable of really being responsible but pretending she is will help her *become* responsible. Below are three quotes that allude to this ‘fake it till you make it’ response. Strawson writes that:

parents and others concerned with the care and upbringing of young children ... are dealing with creatures who are potentially and increasingly capable both of holding, and being objects of, the full range of human and moral attitudes, but are not yet truly capable of either. The treatment of such creatures must therefore represent a kind of compromise, constantly shifting in one direction, between objectivity of attitude and developed human attitudes. Rehearsals insensibly modulate towards true performances (P. F. Strawson, 2008, p. 20).

Recall that, according to Strawson, when we take an objective stance towards someone, we may *pretend* to reason or quarrel with someone. He suggests here that such ‘rehearsals’ will help to render the child capable of true performance. Similarly, Wallace writes:

Because children lack ... [the powers of reflective self-control] ... or are still in the process of acquiring them, it would be unreasonable to hold them fully accountable with respect to the moral obligations we accept. But of course we often treat them *as if* they were accountable agents when they violate those obligations. This is partly because we do not believe even very young children completely lack the powers of reflective self-control (recall that possession of such abilities can be a matter of degree). And partly it is because treating children as if they were responsible is believed to be the most effective way to stimulate the development of their powers of reflective self-control (Wallace, 1996, p. 167).

And Manuel Vargas writes that:

it seems fair to characterize a good deal of parenting and acculturation as bent to the task of expanding the range of circumstances in which the targets have the capacities required for moral responsibility. Initially, much of this happens via feigned attributions of responsibility. In contrast to genuinely holding someone responsible, moral education is typically undertaken in the way characterized by traditional moral influence theorists, that is, with the aspiration of influencing. There is no assumption that the target is a responsible agent. Indeed, the point of feigning praise and blame just is to get children to such a point where they have the capacities that are required for genuine praise and blame (Vargas, 2013b, Chapter 7).

The suggestion is that these feigned responses are common and desirable when it comes to responding to the objectionable conduct of young individuals whom we do not deem sufficiently competent. Could this ‘as if response’ provide a good description of how we do and should respond to the individuals from the three case studies?

Although the ‘fake it till you make it’ response is often mentioned in passing, it has never been thoroughly analysed. What is it exactly that we do when we act ‘*as if*’ the other person is responsible? We may attach consequences that we would typically associate with blaming. But doing so does not amount to a *feigned* response. In order for the response to count as feigned blaming, it seems that we should speak to the person in a similar manner and tone to the one we would use to speak to someone whom we blame, and we thereby *pretend* that we experience blame-related emotions and/or attribute sufficient capacity to the person. We may, for example, pretend to be angry when we discover a toddler is eating the cakes that are on display at a bakery. The sight of this toddler with chocolate all over their face may in fact be endearing and amusing to us, but we *pretend* to be angry because the toddler should learn that he cannot eat the food he encounters in shops.

However, the responses central to the cases we started out with are different. The protagonists do not feign to be affected by the person’s conduct or pretend to attribute ill will and sufficient capacity to the person. Instead, they are truly affected and genuinely address the person. It should also be noted that our responses to children are not always feigned. They are often equally reflective of how we were really affected by the child’s behaviour and characterised by genuine communication with the child about the objectionable conduct.

I suspect that if people *were* to commonly respond with feigned blame to the objectionable conduct of underdeveloped or compromised moral individuals, this would not be very helpful for their moral development. One reason to suspect this is that feigned responses would undermine congruence, which according to a number of psychologists and clinicians is a

prerequisite for a successful therapeutic relationship (Rogers, 2013).² Chapter 7 elaborates on this point. For now it suffices to say that the ‘fake it till you make it’ response does not help to conceptualise and evaluate the responses we are interested in.

To summarise, the responses in the three scenarios 1) are not instances of blaming, 2) do not reduce to excusing or exempting, and 3) do not involve feigned blame. Are there, then, other theories that can account for these responses to objectionable conduct? Hanna Pickard and David Shoemaker have extensively addressed the question of how we do and should hold persons responsible when their agency is underdeveloped or compromised. Do these philosophers provide us with accounts that can help us conceptualise and evaluate the responses at stake here?

2.4 The clinical stance

In a paper called ‘Responsibility without blame: philosophical reflections on clinical practice’, Hanna Pickard aims to make sense of a stance that staff members in a therapeutic community centre where she worked take towards the objectionable conduct of service users with borderline personality disorder. She writes that these service users

regularly behaved in ways that were harmful to staff and other members, even if physical violence was very unusual. They could be emotionally cruel, or extremely angry and threatening without just cause; they might self-harm or disengage from the Community without explanation, provoking high levels of anxiety in others concerned for their well-being; they might shirk their Community tasks and responsibilities, leaving others to pick up the work (Hanna Pickard, 2013, p. 1135).

Pickard notes how the staff considered these service users responsible for such conduct but would not blame them. She concurs that this is the right response. According to her, it is essential that they hold these individuals responsible and accountable in order for them to improve and recover. Blaming would, however, be detrimental to their recovery. According to her:

Blaming service users may trigger feelings of rejection, anger, and self-blame, which bring heightened risk of disengagement from treatment, distrust and breach of the therapeutic alliance, relapse, and, with service users with personality disorder, potentially even self-harm or attempts at suicide: it is essential that compassion and empathy be maintained (Hanna Pickard, 2014, p. 1135).

However, she was also puzzled by this practice of holding individuals responsible without blaming them. How can blaming someone for harm and holding someone responsible for harm

² A therapist is congruent when genuine and authentic towards a patient; this means that their facial and bodily expressions match their words (Seligman, 2006).

come apart? To solve this conundrum, she explains how the blame that is inappropriate is a particular *emotional* response to the person. One may still judge a person to be responsible for harm and even, in a detached sense, judge that a person is blameworthy for harm without also emotionally blaming the person. In order to help mitigate blaming sentiments towards the person, we should take into account the adverse circumstances these service users often grow up in, the stress that they are under, and the stigma they suffer from (Hanna Pickard, 2013, p. 1147).

Does this stance explain how we do and should relate to the individuals from our three case studies? I believe the clinical stance does not describe how we do or should respond in the cases of Earl, Jake, and Anika for two reasons. An elaborate discussion of these reasons is provided in chapter 3. Here I limit myself to an initial characterisation. First, in these cases, detached judgements of blameworthiness are absent and would be inappropriate. Pickard herself believes that individuals with borderline personality disorder or (she argues elsewhere) individuals with addiction have sufficient capacity for choice and control (Hanna Pickard, 2013, 2017a). However, Earl, Jake, and Anika are not considered to have sufficient capacity for blameworthiness; they were not capable of registering the relevant moral reasons in the types of circumstances they were in. Therefore, the judgement that they are blameworthy and responsible for their objectionable conduct would be incorrect.

Second, Louis, Lillian, and Rob do not hold the other person responsible *for* the harm they have done. The expectations or obligations involved are purely forward looking. They are held to have a responsibility to take steps now or in the future because of the harm that has occurred. Attention is drawn to the harm that was done and the impact of their actions. But they are not held responsible *for* the harm. If one were to, for example, hold Jake responsible for the harm he did, one could possibly be justified in asking him to pay for Lillian's therapy sessions, or it could be justified that he undergo some form of punishment. Of course on Pickard's account one would have to hold the person responsible with compassion and without emotional blame. But even so, a call for financial compensation or punishment seems out of place in the case of Jake because he did not disregard Lillian's interests and concerns for the sake of other self-interested reasons. Rather, he could not – under the relevant circumstances – register what Lillian's concerns and interests were.

In the third chapter, which is called 'The Nurturing Stance', I elaborate on these considerations and provide an alternative account of holding responsible without blaming. On the basis of this chapter, one could consider the nurturing stance and the clinical stance to be two competing analyses of how we should relate to individuals whose agency is compromised.

However, I think Pickard's work could be more charitably, and probably more accurately, interpreted as discussing a response that would be desirable in cases that are different from the ones that I discuss in this thesis. I do not provide this particular reading in chapter 3, but I expand on this reading of her work in chapter 7.

In brief, I take it that what the clinical stance tracks is a person's lack of what Pickard and Pierce call 'the will to recover' (Pearce & Pickard, 2010a). The clinical stance is therefore accurate and therapeutically effective in cases where a person does not (yet) have this will to recover. Pickard would, for example, say that an addicted person or a person with borderline personality disorder may behave objectionably *because* they have not yet decided or chosen to change their life around. There can be a number of reasons for this. A person may not trust others to help them do so, they may not have faith that their life can get any better, they may think they do not deserve any better, or they may simply be defiant because others have treated them disrespectfully. These considerations should help to mitigate our blaming sentiments towards the person and help us to have a compassionate attitude (Hanna Pickard 2017, 2013, see also Watson 1993). But we would still attribute to the person a responsiveness to reasons, and these considerations would not take away a person's responsibility for the harm that they have done (Kennett, Vincent, & Snoek, 2015; Hanna Pickard, 2017b).

My account, in comparison, is a response to objectionable conduct that is best explained by insufficient responsiveness to reasons. The question of whether someone has decided or chosen to (learn to) refrain from acting objectionably is moot if the person does not realise that their behaviour is objectionable. When it comes to Earl, Jake, and Anika, the absence of such a decision or choice does not explain why they acted as they did. They may need no encouragement to make this choice once they realise why it is needed. On the contrary, Anika really would not want to act in the way that she did. Jake had no idea that there was even anything objectionable about his actions. Once he realises, he may very well want to improve in the light of those reasons straight away. Their objectionable conduct is primarily explained by the fact that they, under the circumstances, could not have been aware of the relevant moral reason, rather than the absence of a making a wholehearted decision to change. Because this is the agential feature that explains why they acted as they did, another stance is called for. In chapters 3 and 4 I provide this alternative stance and elaborate on the conglomerate of agential features that render it appropriate.

It is an open question whether my account would be (generally) less accurate and therapeutically effective in cases of addiction or borderline personality disorder. Pickard seems to assume that in those cases we are more typically dealing with a lack of will or decision to recover.

But one may contest the evidence she provides for this claim. See, for example, Anke Snoek's research on the role of willpower and self-control in recovery from addiction (Snoek, Levy, & Kennett, 2016). I think that further research is needed in this area to get a clear picture of when and why a lack of willpower and a failure to make the decision to recover obstruct recovery; the research would have to involve extensive feedback from recovered service users (see also chapter 7).

I suspect that the account I develop is especially relevant when it comes to relating to individuals with high-functioning autism, children between the ages of 5 and 10, and persons with an intellectual disability or brain damage.³ It is, of course, not the case that *all* the objectionable conduct of these individuals can be explained by insufficient reason-responsiveness. A person with autism or a child of 10 can be fully responsible for objectionable conduct when they disregard moral reasons rather than fail to register them. A person with autism may, for example, often lack access to moral reasons for acting in a particular way when those reasons present themselves as subtle social cues. And a person with an intellectual disability may easily fail to pick up on reasons in fast-paced scenarios. All I mean to say here is that these types of disabilities or underdevelopments are known to go hand in hand with a lack of access to certain moral considerations in *certain* types of circumstances, and that those individuals are therefore more obviously potential candidates for a nurturing stance than individuals with borderline personality disorder or an addiction.

To conclude, it seems that Pickard is interested in those cases where trying to change a person's will and perspective would be therapeutically effective, because a lack of will to recover primarily explains why the person acted as they did. I am, in contrast, interested in those cases where conduct is primarily explained by an inability to pick up on the relevant reasons for acting in a particular way in certain types of circumstances.⁴ The aim is to develop a stance that recognises this incapacity and a person's agency to correct for this incapacity.

2.5 The tripartite theory

Another account that should be considered in comparison to the account I am about to provide is David Shoemaker's tripartite theory (Shoemaker, 2015). Shoemaker has developed a theory

³ That being said, the nurturing stance can also be an adequate response to the objectionable conduct of neurotypical adults and is part of our repertoire of responses to the objectionable conduct of others in general.

⁴ These explanations may overlap. When, for example, a person with autism is berated for being rude even though she has no access to knowing how her behaviour is rude, the person may end up meeting both agency conditions. She could be insufficiently responsive and unwilling to look into changing her conduct for the sake of others. In these kinds of cases a combination of the stances is called for.

that aims to account for (some) of our responsibility responses to what he calls ‘marginal agents’; among them are individuals with high-functioning autism, an intellectual disability, early stage dementia, and psychopathy. He discusses cases where these particular individuals act objectionably and writes: ‘My own reaction, and the reaction of others who have written about such cases, is a profound unease. This is not, however, the unease of uncertainty. Rather, it is the unease of ambivalence’ (Shoemaker, 2015, p. 3).

He believes this ambivalence can be explained in terms of different ‘responsibility responses’, which track different types of responsibilities. The first class of negative responsibility responses concerns emotions like contempt, disdain, and revulsion. These are responses that evaluate a person’s bad character traits. The second class of negative responses are akin to frustration and annoyance, and these responses evaluate responsibility for bad quality of judgement. The third class is mainly made up of feelings of resentment, indignation, or hurt feelings, which appraise responsibility for moral disregard (Shoemaker, 2015, p. 35). These negative responses can only be appropriate if the person 1) has those character traits, 2) is answerable, e.g. has capacities for judgement, and 3) is accountable, e.g. can have regard for another person’s interests and concerns (Shoemaker, 2015, p. 35). When we relate to marginal agents, our responses are ambivalent because those individuals are not responsible in all three senses of the word (Shoemaker, 2015, p. 123).

For example, according to Shoemaker, in cases of high-functioning autism, we may appropriately evaluate the person’s character or judgements, but we cannot appropriately get angry with the person because autism incapacitates a person’s ability to display regard for the interests and concerns of others (Shoemaker, 2015, Chapter 5).

So in the case of Jake, Shoemaker would suggest that Lillian does and should have an ambivalent response to Jake as a person. She may hold him responsible if he were to make bad judgements, but she cannot be angry with him for having disregarded her interests and concerns because he *cannot* display regard for them. I do not think this account of autism is correct and I discuss why in chapter 5 of this thesis. But the question of relevance in this chapter is whether Shoemaker’s tripartite theory can sufficiently account for Lillian’s response to Jake. It is true that she is not, at the time of addressing Jake, angry with him. But the responses that are central to Shoemaker’s account are all backward-looking evaluations of the person and therefore do not explain how and why Lillian responds to Jake in a way that appeals to his forward-looking responsibilities.

When it comes to the case of Earl, it seems that Louis’ response could qualify as affective disapproval or even irritation. These feelings, on Shoemaker’s account, appraise bad judgement.

They are appropriate if the person has indeed made a bad judgement and can make better judgements (Shoemaker, 2015, Chapter 2). But Earl's brain injury impairs his ability for making rational judgements. On Shoemaker's own account, Earl would therefore lack answerability responsibility. It seems, therefore, that Louis' response would be inappropriate on Shoemaker's account.

But I doubt that affective disapproval is always elicited by an appraisal of bad judgement. I contest that the affective response in Louis' case is more likely to be elicited by the perceived harm that is done to the mother, and would be an accurate response in so far as this perception is correct. These feelings are consequently translated into a response to Earl and an appeal to Earl that does not correspond to any of the three categories that Shoemaker provides.

How would Shoemaker account for Anika's case of manic behaviour? According to Shoemaker, mania can impair all the types of responsibilities (Shoemaker, 2015, p. 123). This may explain why Rob, when he addresses Anika, does not hold it, *in any way*, against Anika that she acted as she did.⁵ However, we are still left with the question of how we should conceptualise and evaluate the way in which Rob *does* respond. The tripartite theory does not provide an answer to this question.

To summarise, the tripartite theory does not provide an account of the responses that are of concern. First, because Shoemaker's responsibility responses appraise a type of responsibility for actions that have happened in the past, whereas the protagonists from our examples only allocate current and forward-looking responsibilities to the person. Second, because his responsibility responses are types of emotions that do not characterize, and are not an essential part of, the responses we are concerned with.

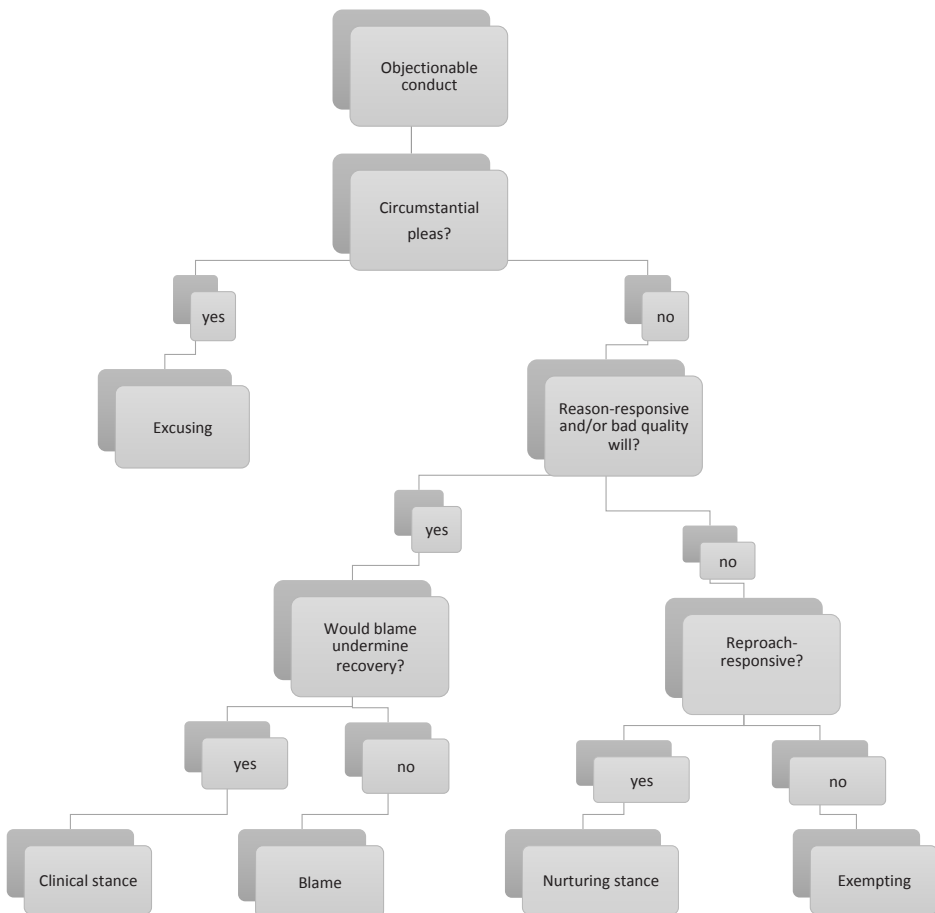
2.6 Conclusion

The chapters that follow will develop the stance that I believe underlies the responses this chapter began with. Below I provide a diagram that indicates where my account will provide a contribution.

This diagram roughly represents the state-of-the-art literature on how we may respond to a person who behaves objectionably. This schematic picture fails, of course, to represent all the nuances of the debate. Some people would, for example, disagree with Pickard and hold that one may blame a person even if this undermines the person's recovery. Furthermore, it is generally

⁵ One may, however, feel anger when someone behaves badly as a result of mania and at the same time understand that the person is not responsible. It is unclear whether Shoemaker's account can explain why we can experience the emotions he discusses when we know the person is not responsible in a way that would justify this emotion.

thought that blame would still be wrong if the person who is blaming does not have standing to blame and that further conditions have to be met for it to be appropriate to *express* blame to a blamee. Sceptics' objection to all of this would be that blaming is *never* desirable. For the sake of clarity, I have not represented all these disagreements in the diagram. In this thesis I bracket the debate about whether blaming is ever appropriate, and if so when. Instead I focus on how we do and should relate to a particular group of individuals who would not be considered blameworthy on both compatibilist and sceptic accounts of moral responsibility. In green below I indicate where I plan to contribute to the field.



In the next three chapters I provide an account of the nurturing stance and the agency condition that renders this stance an appropriate response (reproach-responsive in the diagram above) and discuss the theoretical implications of a number of accounts of responsibility. In the last two chapters I discuss the practical application of a nurturing stance in, respectively, the contexts of criminal justice and psychiatry.

3 The Nurturing Stance: Making sense of Responsibility without Blame

Abstract

Mental health-care clinicians report that they hold patients responsible for morally objectionable behaviour but at the same time consider blaming attitudes to be inappropriate. These practices present a conundrum for all Strawsonian theories of responsibility. In response to this conundrum, Pickard has proposed severing the Strawsonian connection between being responsible and being an appropriate target of blaming attitudes. In this chapter I will argue that her solution fails to explain the practices at stake and provide an alternative solution that uncovers an under-theorized stance we take towards those whose abilities are underdeveloped or compromised.

Keywords: responsibility, agency, exemption, blame, Strawson, reactive attitudes, psychiatry

The central aim of this chapter is to make sense of a type of responsibility practice that is central to mental health-care contexts. This practice can be characterized as ‘allocating responsibility without blame’ (Hanna Pickard, 2013). When a patient behaves in a manner that does not meet moral standards – e.g. verbally abuses someone, throws things across the room or is aggressive and threatening towards other patients or staff – it is considered good practice to hold the patient responsible for this, but blaming the patient is typically considered inappropriate. It may at first seem difficult for both clinicians and philosophers to make sense of this practice. This difficulty is due to the intuitive link between someone’s responsibility for harm and the appropriateness of blaming attitudes towards this person. This link is one that has been stressed most prominently by Strawsonian theories of responsibility. In the first section I will expand on the apparent tension between Strawsonian theories and these practices of allocating responsibility without blame.

There is one rather straightforward explanation for refraining from expressing blaming attitudes to a patient in psychiatric practice that does not pose a problem for Strawsonians. The expression of blame can be detrimental to the therapeutic relationship and to the patient’s wellbeing. This provides reasons not to blame the patient, especially for a clinician whose main concern is the patient’s recovery. However, these reasons do not undermine the appropriateness of blaming attitudes; they only render the expression of these attitudes inappropriate in a clinical context. I will briefly discuss this non-threatening explanation in section 2. But, although this may be an accurate explanation for some practices of holding patients responsible without blaming them, it certainly does not adequately explain all of these practices.

Sometimes clinicians refrain from blaming a patient for causing harm or abuse because the patient is not considered blameworthy. Even in these cases, however, the patient is, in a sense, held responsible. The challenge this presents us with is a more serious one. It has to be explained how and why a patient is held responsible after behaving in a morally problematic way when he is, at the same time, not considered to blame for this behaviour and blaming attitudes are *therefore* inappropriate. The rest of the chapter is concerned with solving this challenge.

Hanna Pickard proposes solving this challenge by severing the link between someone’s responsibility on the one hand and the appropriateness of moral blame on the other hand. According to her, one’s responsibility should be based on the normative capacities that one has, but responsibility can be *detached* from blaming attitudes (Hanna Pickard, 2013). In section 3 I will argue that Pickard’s solution fails to explain the practices at stake. I argue that she in fact fails to detach the appropriateness of blame from the normative capacities that one has to have to meet moral expectations.

As an alternative solution, I propose that responsibility without blame does *not* track full-blown normative capacities and responsibility for harm, but instead has underdeveloped or compromised capacities as its target. Underdeveloped or compromised capacity grounds the type of responsibility that is central to responsibility without blame: a responsibility to work towards developing or repairing one's own moral abilities. I dub this form of holding responsible 'the nurturing stance'. This responsibility to engage with one's own impairments differs from but closely relates to traditional responsibility for harm.

In conclusion I argue that the nurturing stance does not undermine the Strawsonian theories at stake but does provide an important addition to them. In practices of allocating responsibility without blame, patients are held to the responsibilities they have to become able to meet moral expectations. This stance is a common but under-discussed aspect of our moral practices. The Strawsonian typically overlooks it by suggesting that we may appropriately respond to a moral transgression by *either* blaming *or* exempting the person; this stance shows that such a binary approach is misleading. People who transgress a norm are sometimes not blamed for this because they couldn't help transgressing, but when they can become able to meet our norms, our attitude towards them in response to this transgression is not properly described as exempting.

3.1. Strawsonians and the Clinical Conundrum

P.F. Strawson observes that there is a connection between our reactive attitudes and being a responsible agent. Strawson notices that when we respond with resentment, indignation or similar blame-related attitudes to another person's actions, we typically consider this appropriate if the agent disregarded our shared moral expectations and demands. But if someone's abilities to meet these interpersonal demands are severely underdeveloped or impaired, we generally feel differently: we modify, suspend or do not even experience these attitudes, or, at the very least, think we *should* not respond with indignation or resentment (Peter F. Strawson, 1962). Blaming attitudes are considered to be an inappropriate response to someone whose underdeveloped or impaired normative capacities render them unable to meet the demand at stake. This does not mean that we never resent someone whose transgression was due to impaired normative capacities, but it is generally accepted that we *should* not do so. Blaming the incompetent for transgressing a norm that they cannot meet is generally considered inappropriate.

Many philosophers after Strawson have agreed that being responsible implies being an appropriate target for praising or blaming attitudes. Some have argued that these attitudes generally indicate or track the criteria that make for a responsible agent (Wallace, 1996). Others

prefer to say that our disposition to praise and blame is what constitutes the responsibility of the subject of such praise or blame (Shoemaker, 2015; Watson, 2004; for discussion see Todd, 2016). Which criteria exactly are singled out by these attitudes and hence are the responsibility-making features of an agent is also a matter of debate. But although the details of this relation are fleshed out differently, many maintain that an agent is responsible if, and only if, she is an appropriate target of blaming or praising attitudes (i.e. Fischer & Ravizza, 1998; Shoemaker, 2015; Wallace, 1996; Watson, Gary, 1993).

But this connection between blaming responses and responsibility gives rise to a clinical conundrum. The clinical conundrum arises if one takes a close look at nursing practices in psychiatry. Clinicians hold patients to be *responsible* for their actions when they misbehave, but often consider blaming them to be inappropriate (Hanna Pickard, 2013). My understanding is that this practice should be imagined as follows: imagine that a patient is throwing a chair across a living room, thereby breaking things and endangering and upsetting other patients. In response to this transgression, good practice prescribes that a clinician speaks to the patient later and says something like ‘I understand that you were upset, but this behaviour cannot be accepted here. What are you going to do about it? And how can we help you?’ Clinicians report that they consider the patient to be responsible for his own behaviour and hold him responsible for it in such a case (Brandenburg, unpublished). That the patient is held responsible in this practice is *prima facie* a plausible description of what is going on here. I take it that the patient is being held responsible when it is pointed out to this person that his or her behaviour was problematic and it is made explicit that changing this is expected of her and (at least partly) considered up to him or her. This is the form of holding responsible that will be central to this chapter.

Note that the patient is not blamed in such a situation. The clinician would not address the patient with anger, resentment or indignation, nor would she say something like ‘You had no concern for us or others’ or ‘You have been negligent. You should know better than this’ or ‘This is your fault!’ and so on. Such responses would be considered grossly inappropriate.

Pickard describes how this practice at first struck her as a philosophical and clinical conundrum. She could make sense of the idea that, despite appearance, patients are not responsible or culpable because of their disorder and hence are not to be met with anger or resentment. She could also make sense of the idea that despite their disorder they are responsible and culpable and hence liable to anger or resentment. But she found it difficult to make sense of this stance of allocating responsibility without blame (Hanna Pickard, 2013, p. 1135). Family, friends and clinicians of patients who transgress in these ways and often encounter the same difficulty in making sense of this practice of allocating responsibility without blame. The two

questions that will be central to this chapter are as follows: how can this conundrum be solved, and how is the Strawsonian relation between reactive attitudes and responsibility affected by this solution?

3.2. Responsibility without Expressed Blame

Surely some of the practices described above may concern patients who *are* blameworthy for what they have done. Whatever the exact criteria for blameworthiness are, a patient may – just like anyone else – meet these criteria on some occasions (M. McKenna & Kozuch, Benjamin, 2015; Sripada, Chandra, 2015). Maybe the patient throwing the chair across the room knew very well that he had got himself into a situation that would make him uncontrollably upset, and maybe it would also have been quite easy for him to avoid getting into this situation. Even then, responding in a blaming fashion is typically considered inappropriate within a clinical context.

A blaming response can undermine the patient's trust in the clinician and make him feel rejected, which in turn compromises the therapeutic relationship and the patient's recovery. A clinician may also find it inappropriate to blame the patient because such a response doesn't fit well with her professional and caring role. Pickard seems to suggest that it is *always* crucial that clinicians refrain from expressing blaming attitudes, even when a patient is blameworthy, so as not to risk damaging therapeutic efficacy (Lacey & Pickard, 2013a; Hanna Pickard, 2013). I believe that this is taking it too far. There are reports of cases in clinical settings in which a temporary and moderate expression of anger or indignation has been conducive to a better therapeutic relationship and has actually worked towards certain therapeutic aims (Brandenburg, unpublished). But blaming attitudes are better held at bay if they would stand in the way of a good therapeutic relationship and meeting therapeutic aims even when a patient culpably misbehaves.

Of course, more generally speaking, giving expression to blaming attitudes is not always the appropriate thing to do, even in cases where the person is blameworthy for transgressing a certain norm. Angela Smith has before pointed out how these observations would pose a serious objection to those Strawsonian theories that suggest that the expression of blaming attitudes is *always* fair and appropriate when a competent agent expresses ill will or insufficient regard (i.e. is blameworthy)(Wallace, 1996, pp. 187–193; A. M. Smith, 2007).

But the existence of further conditions for the appropriate expression of blame ultimately does not challenge the Strawsonian relation between the appropriateness of blaming attitudes and someone's responsibility for harm. If the subject from our example is blameworthy, blame is still an appropriate response in a similar sense to the way in which fear would be an

appropriate response to something being dangerous (Russell, 1992). This appropriateness relation still holds even if other considerations trump the appropriateness of *expressing* blame in a specific context. A comparable example is that although it may be very unwise and ineffective to give expression to fear when confronted with a roaring lion, the experience of fear is not thereby inappropriate. On the contrary, as an appraisal of the situation this is an accurate response.

In our imagined case, where the patient is blameworthy for throwing a chair through the room, the *experience* of blame-related attitudes in response to this action would also still be appropriate. Whilst expressing reproach may not be efficient or useful in this case, it is certainly apt to feel reproachful in the light of what the patient has done. In addition, this solution to the conundrum, strictly speaking, does not provide us with a case of responsibility without blame; it only provides us with a case of responsibility without *expressed* blame.

But this solution does not apply to all those cases in which a clinician holds a patient responsible without blaming the patient. In some cases blaming attitudes *would* be inappropriate because it would be wrong to say that the patient *is* blameworthy for transgressing a certain norm. Imagine that the patient in our case could not foresee that getting into this situation would get him as upset as it did, and that the emotional overload was just beyond his control at that very moment. He felt unsafe, got very stressed and threw a chair across the room, which at that point seemed for him to be the only way to deal with his feelings. We also know that because of his mental illness he is much more likely to get stressed in such situations and never really learned how to cope with that stress. Even in such a case, and in many similar cases, a patient will typically be held responsible for this behaviour in the sense described in the first section. How are we to explain that the clinician holds the patient responsible when the patient is *not* blameworthy for her behaviour? In section three I will discuss Hannah Pickard's solution to this challenge and argue that it fails to explain the practices at stake.

3.3. Responsibility without Blameworthiness

Pickard also points out that psychiatric patients often cannot be considered blameworthy for what they do. In order to explain why patients are nonetheless held responsible, she proposes separating responsibility for harm from blameworthiness for harm. According to her, patients *are* responsible for norm-transgressions but are not blameworthy (Hanna Pickard, 2013, pp. 1141–1142). Hereby Pickard provides a solution to our challenge. The patient is held responsible for the transgression because he *is* responsible for the transgression. But blaming the person is deeply inappropriate, because the patient is not blameworthy for the transgression. This solution

drives a wedge between being responsible on the one hand and being an appropriate target for blaming attitudes on the other. But her solution is conceptually puzzling and not satisfactory as it stands.

Pickard defends her conceptual solution of ‘responsibility without blame’ by probing the reader’s intuitions. As readers we are asked to compare the following:

‘1. Service users may be responsible for verbal aggression towards clinicians but not blameworthy, because they are acting to relieve high levels of psychological distress, and lack alternative coping mechanisms.’ (Hanna Pickard, 2013, p. 1140)

‘2. *Service users may be morally responsible for verbal aggression towards clinicians but not blameworthy, because they are acting to relieve high levels of psychological distress, and lack alternative coping mechanisms.’ (Hanna Pickard, 2013, p. 1140)

The distinction between the two descriptions is demonstrated by the fact that 1 is intuitively plausible whilst 2 is not, and it explains, according to Pickard, how it is ‘possible to be responsible and treated thus, for actions which are morally wrong but for which one is not blameworthy because one has an excuse’ (Hanna Pickard, 2013, p. 1142). But it is not very clear why our intuitions should diverge here if the actions at stake are indeed morally problematic and the agent is responsible for those actions. As far as my own intuitions are concerned, if the patient *has* capacities for control then intuitively he is both responsible and blameworthy, and if he does not have the capacities for control, then he is neither responsible nor blameworthy. The distinction between responsibility and moral responsibility doesn’t help avoid this conclusion.

Pickard is explicit in sentence 1 about explaining patients’ responsibility on the basis of their ‘capacities for choice and control’. According to Pickard, it is because patients *have* such normative capacities to refrain from inflicting abuse or harm that they should be held to account for their actions. But if it were true that patients have these capacities, then it would be intuitive to say that they are accountable *and blameworthy* for the things they do. They are culpable because they have the capacities for choice and control and hence *can* control their behaviour and can choose to refrain from doing harm. If Pickard is right about this, we should simply say that patients are responsible and blameworthy for the harms and wrongs they do but that one should not respond by means of expressing blaming attitudes because this is detrimental to therapy (Hanna Pickard, 2013, p. 1142).

The second half of statement 1 raises even more questions. According to Pickard, patients are often not blameworthy because they are excused. The patients are excused because ‘they are acting to relieve high levels of psychological distress, and *lack* alternative coping mechanisms’ (Hanna Pickard, 2013, p. 1140). This does indeed seem to be a likely explanation for the absence of blameworthiness, or at least for a strong mitigation of blameworthiness. But now it is unclear why Pickard wants to maintain that patients *have* the capacities for choice and control *and* are therefore responsible for verbal aggression. If one has problems with regulating emotions and lacks alternative coping mechanisms, one is – all other things being equal – neither responsible nor culpable for the harm one does, *because* one’s capacity for control is impeded. It seems strange to hold someone responsible for a transgression like throwing a chair across a room or verbal aggression when this transgression is due to a serious control impediment.

The one way in which it may make sense to say that the person is nonetheless responsible in such a case is when one means to say that the persons is *causally* responsible for what happened. But note that this can’t be right either, because if one is merely causally responsible it is mysterious why one would be *held* responsible for what happened. Imagine someone is pushed over and breaks a vase as they fall; they are therefore causally responsible for breaking the vase. It would in such a case be totally ludicrous to go up to this person and tell her that this behaviour cannot be tolerated and that she should do something about it. Hence, we are still left with the fascinating question of why these patients in practice *are* more than just causally held responsible if they are not culpable because of their impaired ability for control.

An interesting answer to this question remains obscured because of ambiguity about what it means to have capacities for control. I think Pickard is wrong to say that the patients in our examples simply ‘have the capacity for choice and control’ that is required, because these patients cannot control themselves in certain stressful situations. But there is a sense in which it also wrong to say that these patients *do not* have the capacity to control themselves in such situations. I will explain why, and this will in turn provide us with a clue for answering the challenge at stake.

3.4. Ambiguous Capacities

In this section I will argue that there is an important sense in which patients who are held responsible without blame *do not* have the capacities to meet certain expectations to refrain from causing harm or abuse but another sense in which they *do* have these capacities. The crux is that in natural language, when we claim that someone has the capacity to do something, we may also be referring to capacities that are underdeveloped or compromised.

Compare the following sentences:

1. You have the capacity to walk
2. You have the capacity to be a great leader

Or:

1. You have the capacity to resist the extra glass of beer
2. You have the capacity to get your driving licence

Note that the meaning of ‘capacity’ is often different in sentences of type 1 and type 2. When put next to each other this becomes quite clear. The word ‘can’ or ‘capacity’ in type 1 sentences refers to something someone is normally able to do. Human beings can usually walk and they can resist the extra glass of beer in most circumstances too. In these sentences the word capacity simply refers to abilities that we take the average human being to have. When I say that you have the capacity to walk, I mean that you can get up and do so *right now*.

But being a great leader requires experience and learning. Most people can be said to have the capacity to be a great leader because they *can* be a great leader with sufficient practice and experience. And, even when you have the skills to be a great leader, there may be periods when your leadership falters because, say, lack of sleep or stress compromise your ability to exercise good leadership. However, even when your capacity to be a great leader is underdeveloped or compromised, I would not be lying when I say to you that you have the capacity to be a great leader. Similarly, when someone tells me that I have the *capacity* to get my driving licence after I have just had a driving lesson that left me in utter despair, this does not mean that I can get into a car and pass the driving test *right now*. On the contrary, it may very well be fatal to me, or others, if I try to do so. But still, this person is not necessarily wrong by telling me that I *can* get my driving licence, because with sufficient practice and training even I will probably be able to pass the test and get my licence.

The use of the word capacity in type 2 sentences refers to abilities that are available to us in the future after we have engaged in learning processes and/or have overcome the obstacles that are needed to develop or restore these abilities. These capacities differ from those that are developed and uncompromised. From now on I will refer to the latter as actual capacities, in contrast to underdeveloped or compromised capacities. Now that these different meanings of the word capacity are clear, I can return to Pickard’s claim that patients have the requisite capacities for meeting the demands at stake. Pickard, for example, wants to say that

- Patients have the capacity to refrain from engaging in alcohol abuse; and
- Patients have the capacity to refrain from engaging in verbal aggression

My contention is that in psychiatry, these two sentences are typically meant to refer to capacities that are either underdeveloped or compromised. One possible explanation for the patient's failure to meet an expectation or standard at this moment or in specific types of circumstances is that her ability to meet such expectations is underdeveloped or compromised. I do not want to claim that this *is* what Pickard means. I'm not sure it is, but I do think it is one good way to make sense of what she may mean. Though the average adult has the capacity to refrain from, say, using verbal aggression or abusing alcohol, many patients in the clinic have this capacity in an underdeveloped or compromised sense. Their abilities to, for example, refrain from being verbally aggressive or engaging in other types of abusive behaviour are yet to be developed or restored by, for instance, acquiring certain habits, lowering their levels of stress or anxiety and/or learning how to cope with their emotions. The transgressions from our examples are often due to excessive anxiety, stress or other forms of emotional flooding. One further remark is therefore required here. It has been argued that being overcome by stress or anxiety may *fleetingly* mask an actual capacity (M. McKenna & Kozuch, Benjamin, 2015, p. 94). It should be noted that in clinical settings transgressions are more often due to robust forms of emotional flooding and hence to inhibitions of actual capacity rather than to 'incidental masking' of capacity to which everybody may be subject at times.

Norm transgressions in psychiatry are often, though not always, due to the patient's underdeveloped or compromised ability to refrain from causing harm or abuse. Pickard would be wrong to say that these patients have the capacity for controlling and voluntarily choosing to live up to these expectations if she means to say that they have the actual ability to do so to the same extent that the average adult in the street. But they do have the capacity in an underdeveloped or compromised sense. Capacity is an ambiguous term that can, among other things, refer to abilities that are underdeveloped or compromised (Morriss, 2002, pp. 52–60). Making this explicit will help us to explain why patients are at times held responsible for their actions even though they are not, or are hardly, blameworthy for what they have done.

3.5 The Nurturing Stance

Recall that the challenge is to explain how it is possible that patients are held responsible after a norm-transgression but are at the same time not blameworthy for this norm-transgression. My

solution to the challenge is the following explanation: although actual capacities are typically coupled to both responsibility and culpability for a moral transgression, underdeveloped or compromised capacities are coupled to another type of responsibility. The ability to develop or restore certain abilities gives rise to the responsibility to engage with one's own failure by means of developing or restoring these abilities. This responsibility becomes salient and is something one is held to when a reasonably grave norm transgression takes place. Strictly speaking, the patients are then not responsible for what they have done, but they are responsible for altering these 'types of doings'. To see how actual capacity, compromised capacity and the absence of capacity give rise to different responsibility practices, consider the following normative scenario.

On a forgotten archipelago there is a community that has little food resources except for the fish swimming at the bottom of the sea. An average member of this community has the skills to dive for the fish and feed herself and some of the young, elderly and sick. But catching fish is hard, and there are times when there are hardly enough fish to live from. It is of course therefore expected that all the able community members dive for fish, and those who are able to do so but don't are considered responsible and culpable for this. Other members of this community may reasonably get angry with these people, and surely they don't deserve to get any fish.

But there are also members of this community who do not learn to dive and hold their breath long enough as quickly and easily as the others because of their fear of the deep, dark water. One of these members is Toddy. One often finds him standing hesitantly on the shore for a while and once he gets into the water he doesn't manage to dive all the way to the bottom. It seems inappropriate to respond with resentment or indignation to his failure to get himself some fish because it is not that Toddy *disregards* the norms of his community; his ability to live up to the norms is underdeveloped and compromised by fear.

But the other members of this community will probably not cater to his need for fish as they would for the sick, the elderly and infants. Though they may give Toddy some fish, this would probably be on the condition that he does something about overcoming his fear and further develops his diving skills. Toddy is not like the elderly, the sick and infants, who cannot be expected to dive for fish *at all*. Toddy can't do it *now* but there is reason to assume that he may become able to fend for himself and, given the scarcity of fish and the work involved in diving for it, Toddy will be expected to do so.

In terms of the interpersonal practices of the community, this expectation is made explicit when Toddy fails to get fish again. Maybe not every time, but most of the times when Toddy comes out of the water empty-handed are occasions for the others to remind him how

pressing the community demand to fend for your own fish is. They will perhaps encouragingly tell him, 'You have the capacity to do this, Toddy, we know you do!' or say more serious things to him, such as, 'You know things can't go on like this. We have many mouths to feed.' This is surely a form of holding Toddy responsible for this type of behaviour. But it is not a form of holding him responsible and culpable for his specific past failure to meet a norm. That would only be appropriate if Toddy already had the abilities to live up to this norm. Toddy, however, does not get exempted like the infants, the elderly and the sick, because he does have the abilities to meet this norm in an underdeveloped or compromised sense. These abilities ground a form of holding responsible that targets and bootstraps his potential to develop or repair them.

Similarly, if in our own lifeworlds climbing a mountain were to somehow become a strong moral requirement, we may ask everyone to learn to climb a mountain because they 'can'. Our underdeveloped or compromised abilities imbue us with certain responsibilities when these abilities are or become related to strong normative demands. In the process of acquiring these capacities, we typically *keep* holding one another accountable by reminding each other of this important expectation and by making explicit that we need to live up to it on those occasions that we fail to do so. But note also that we thereby do not *blame* one another for failing to climb a mountain *now*. Instead, when someone fails to climb a mountain we remind this person that she should *become* able to do so and that it is required of her to keep trying and learning. We are more likely to show sympathy and understanding in response to failure. This goes hand in hand with reminding someone of and holding someone to their responsibilities. In such scenarios we typically hold one another responsible without considering the other to be culpable and without responding (experiencing or expressing) with feelings of resentment, contempt or agent-directed anger.

The scenarios just discussed involve similar elements to what happens in those situations within a psychiatric setting where a patient is *not* considered to be blameworthy for moral harm but is held to account in response to this transgression. Just like Toddy's failure to meet an expectation, the transgressions in clinical settings may be due to underdeveloped or compromised capacities. In a clinical setting, one may similarly be held responsible by being urged to work towards acquiring the sort of control that enables one to refrain from important norm-transgressions. And, of course, where possible the clinicians will help and assist in this learning process.

Such a response is illustrative of good practice surrounding norm-transgressions that come about as a result of stress, troubles with coping or other things that stand in the way of the ability to meet certain expectations. This response appropriately recognizes and targets the

underdeveloped or compromised capacities the patient has, is conducive of therapeutic effectiveness and is in line with duties of care. I believe that in such a scenario, the patient is then indeed held responsible because he *can* do otherwise. But he is not blamed because he *can't* do otherwise. Both are true because *can* here refers to different *types* of capacities.

Please note that this does not mean that the patient in question *was* already able to learn not to transgress norms and *should* have already done so. That would reduce my account to a form of indirect responsibility, which would permit blame. This is atypical for patients in psychiatry. It is more plausible to assume that they are already engaging in this learning process, do not know how to engage in a learning process, do not believe or trust themselves to be able to do so and/or can only do so with a sufficient level of support that they did not have before. In these cases, blame would be an inapt response, but there is some value in holding patients to an expectation by urging them to engage in developing and repairing their abilities. Such a nurturing stance bootstraps a patients' agency to develop or repair capacity and recognizes them as people who are able to do so.

Before discussing the implications of my solution for Strawsonian theories, I want to end with some reservations. Firstly, in practice it is of course very difficult to determine whether normative capacities are compromised or underdeveloped. By providing such a schematic overview, I do not want to deny the difficulties involved in assessing whether norm-transgressions are due to compromised or underdeveloped capacities or to insufficient regard or ill will. Secondly, it should also be noted that the capacities or incapacities at stake are context-dependent and norm-dependent. They are norm-dependent because the capacity to meet one norm can be independent of the capacity to meet another norm. One may, for example, have difficulties in refraining from abusing alcohol whilst being perfectly able to refrain from using verbal aggression, or the other way around. These capacities are also context-dependent because one may be more or less able to do something in different contexts and at different moments in time (Vargas, 2013c). This further complicates real-life assessment.

These assessment problems, of course, also apply to assessing whether someone has the abilities to develop or repair his or her underdeveloped or compromised agency. In reality one is often uncertain about these abilities when relating to someone who struggles to live up to moral expectations. But where uncertainty about abilities gives one reason to be careful and hesitant in allocating blame for harm, the uncertainty about one's abilities for learning does not, in the same way, give one reason to be careful in adopting a nurturing stance towards the person.

It is often more respectful and constructive to persist for quite a while in assuming and hoping that the addressee has abilities for learning and improving that can be engaged and bootstrapped by a nurturing stance, even when this is uncertain. Note how people sometimes engage abilities for learning that are quite probably themselves underdeveloped. Carers can be seen to respond to children as young as two or three in a nurturing manner: they tell them what they ought and ought not do in response to their objectionable behaviour. It is at quite an early age already that a child may be minimally responsive to this stance and may be engaged in his or her own development in some minimal way. As they grow older, this ability for engagement further increases and develops. The same can be true for the abilities of psychiatric patients.

But in cases of uncertainty about the (level of) ability to engage in developing and repairing one's agency, a continuous failure and lack of improvement on the part of the addressee of the nurturing stance provides more reason to wonder whether one is expecting too much of the person than reason to allocate blame for these failures. Uncertainty about someone's (level of) abilities for learning does not undermine the appropriateness of employing a nurturing stance towards this person but does count against blaming him or her for subsequent failures to learn and improve. I take it that this is why we generally are –and should be– hesitant to blame children until they are in quite a late stage of their development and avoid blaming patients in relapse too.

3.6 Strawsonians and the Nurturing Stance

The nurturing stance does not undermine the Strawsonian story about the appropriateness of blaming attitudes, but it does provide an interesting addition to them. In order to meet Pickard's challenge, Strawsonians need to explain why the patient is not blameworthy for a transgression but is nonetheless responsible without thereby denying that 'being responsible' means 'being an appropriate target of praising and blaming attitudes'. I do not have the space to discuss all Strawsonian accounts of the required abilities for being responsible and an appropriate target of blaming attitudes here. My only aim in this section is to show how prominent Strawsonian accounts can perfectly explain the following:

1. Why the patient in our example is not an appropriate target of blaming attitudes *because* the patient lacks the abilities required for responsible agency, i.e. is not responsible, and
2. Why the patient *is* – in another sense – responsible *and* that this responsibility is still connected to being an appropriate target for praising and blaming attitudes.

Fischer and Ravizza argue that the abilities required for being an apt target of praising and blaming attitudes are at least the ability to grasp and apply moral reasons and the ability to act in the light of such reasons, and Wallace defends a similar position (Fischer & Ravizza, 1998; Wallace, 1996). The patient in our example may very well be able to grasp moral reasons. He knows, for example, that the risk of injury and damage provides a reason to refrain from throwing a chair across the room. But he *lacks* the ability to act in the light of such reasons in stressful situations. Because his ability is impeded in this sense, blaming him for transgressions within such contexts would be inapt because in *these contexts* some of the abilities required for responsible agency are impaired. Wallace and, Fischer and Ravizza can also explain why, as I have argued, the patient bears some responsibility for *becoming* able to handle stressful situations. The patient has this responsibility because he can grasp and understand why the norm ‘you should become able to refrain from throwing things around when you get stressed’ is a legitimate one, and he can – possibly with some help – come to act in the light of this norm, i.e. there is reason to assume that he *can* become able to refrain from throwing things around. But note that this also makes him liable to blame again. Those features that make him responsible are also the features that make him an apt target for blame if he does not comply with the demand to improve himself out of insufficient regard for this norm. The Strawsonian connection is still intact.

Other theorists, like Watson and McKenna, focus on the abilities required for being morally *addressed* (1998; 1993). According to such theories, the patient in our example may at first seem to be responsible and blameworthy to the extent that he can be *addressed* in terms of reactive attitudes. He can, in other words, understand what is being communicated by such attitudes. But the ability to understand praising and blaming attitudes is not enough on these accounts. Watson writes that the demand expressed by these attitudes should be inhibited in response to the person’s incompetence ‘in some or all respects for “ordinary adult interpersonal relationships”’ (Watson, 1993, p. 123). According to Watson, if a person cannot express himself because of great strain or stress, the demand that he should express good will and regard should be inhibited. The person is not then appropriately addressed by reactive attitudes that communicate a demand the person could not meet because of such impairments (Watson, 1993, p. 131). McKenna similarly argues that apt targets of our responsibility practices should understand what others communicate to them within those practices *and* should not be impaired in such a way that they cannot themselves –through their deeds- communicate within those practices (M. McKenna, 2012). Because of this, both McKenna and Watson *are* able to explain

why the patient in our example is *not* an appropriate target for blame but is still responsible in another sense.

The patient is not an apt target because in the specific types of contexts in which the transgression takes place (stressful contexts), he is not able to communicate within a normative practice; he cannot – within such contexts – respond to the demand that would be expressed through reactive attitudes such as anger or indignation. Therefore the patient is not an apt target for blame when it comes to transgressions that occur under stressful circumstances. But when it comes to another demand that is central to our practices, the patient *can* understand, communicate and participate. The demand ‘to develop the skills necessary for norm compliance in stressful contexts’ is a norm that the patient can understand when it is being communicated to him *and* he can ‘communicatively participate’ within *this* aspect of our normative practice, i.e. he can respond to this demand. This is then the kind of responsibility he does have and this responsibility is *still* something that would make him liable to be a target of blaming attitudes if he were to disregard the demand for development. Hence, the conundrum, as I understand it, ultimately does not seriously undermine the Strawsonian connection on these prominent accounts.

But the nurturing stance does provide an important addition to a Strawsonian theory. Most Strawsonians tell an incomplete story when it comes to our repertoire of responses to norm transgressions. The suggestion often seems to be that we can *either* evaluate the person as responsible for the transgression, in which case blaming attitudes are an appropriate response, *or* we can consider the person to be exempted from responsibility for the transgression, which means that our blaming attitudes are suspended and the person is seen as someone to be ‘managed’ ‘handled’ or ‘controlled’ rather than to be held responsible and reasoned with (Peter F. Strawson, 1962; Wallace, 1996; Watson, 1993). Their relative silence about other forms of responding to transgressions suggests there are none, and this gives rise to the conundrum this chapter began with.

The nurturing stance is a response to a transgression that crucially differs from this description of interpersonal exempting. When a person is only exempted in response to harm that has been done and is in no way held responsible, we consider the person to be unable to develop or repair the inabilities that explain why a norm was transgressed. We, as it were, ‘give up’ on seeing the person as a current *or potential* participant in our norm-guided moral practices and instead see the person as someone to be ‘managed’ or ‘handled’. But in the case of a nurturing stance, the harm done gives rise to a future responsibility that is expressed in the way that we relate to the person after he non-culpably fails to meet an interpersonal demand. We

urge the person to *take* (or keep taking) responsibility for developing and repairing his own compromised or underdeveloped abilities. When one is subject to the nurturing stance, one is reasoned with and called upon. By adopting a nurturing stance, one recognizes and respects the other person as someone who can come to meet the shared expectations and norms at stake and bootstraps these capacities. By (only) exempting the person, these abilities for development or repair would not be duly respected, and by blaming the person we would misrecognize one's underdeveloped or compromised agency. The nurturing stance is an interpersonal response to norm-transgressions that does not reduce to interpersonal exempting or to interpersonal blaming.

I suspect that in all interactive processes surrounding the development or repair of an ability to live up to a norm, this form of relating to one another can be found. Because the nurturing stance is so common to our interpersonal practices and because it is different in an important way from both exempting and blaming a person, it provides an important addition to any Strawsonian theory of responsibility.

3.7 Conclusion

I have argued that the Strawsonian connection between 'being responsible' and 'being an appropriate target of praising and blaming attitudes' seems to give rise to a conundrum in mental health-care practice. The conundrum is that patients are often not blamed but are held responsible after norm-transgressions occur. The most interesting version of this challenge is the one in which the patient is not blamed because he is not blameworthy, but nonetheless is held responsible. The challenge was to explain why patients are held responsible when they are not blameworthy and to spell out what this implies for the Strawsonian theories at stake. I have argued that these patients are held responsible because they have an underdeveloped or compromised capacity – rather than an actual capacity – to meet the norms that they transgressed. They are therefore held to their responsibility to engage in a process of developing or repairing these capacities. This solution to the conundrum ultimately does not undermine the Strawsonian theories at stake. But it brings to the fore a type of response to norm-transgressions that is a central but under-theorized aspect of our interpersonal practices.

4 Inadequate Agency and Appropriate Anger

Abstract

Communication and cultivation accounts of responsibility (CC accounts) argue that blaming has an important communicative and agency-cultivating function when addressed at someone we consider to be deserving of blame. On these accounts, responsible agents are agents who can understand negative reactive attitudes and are (generally) sensitive to their moral-agency cultivating function.

In this chapter I examine our reproachful engagements with agents whose moral agency is underdeveloped or compromised. I discuss how these engagements compare to blaming on CC accounts and argue reproachful engagements can have an important communicative and agency-cultivating point.

I then go on to propose an addition to CC-accounts that explains how underdeveloped and compromised agents can be held responsible. I will show how this addition resolves ambiguities in accounts from McKenna, Vargas and McGeer. I conclude the chapter by anticipating an objection CC accounts could raise: that reproach and blame are co-extensive.

Our emotional engagements with compromised or underdeveloped agents are a central but under-theorized aspect of our interpersonal practices. Children and service users in mental health care are at times –temporarily or permanently- insufficiently able to comply with moral norms, and their carers then do not consider them blameworthy for transgressing them. Nonetheless their behaviour may give rise to negative emotions and experts consider it at times appropriate to reproach the service user or child (Brandenburg, unpublished interview data). But in philosophy it is commonly assumed that negative reactive attitudes are inappropriate and best suspended if their addressee lacked ability to comply with moral norms (M. McKenna, 2012; Vargas, 2013a; Wallace, 1996). This chapter aims to resolve this apparent tension between practice and philosophical theory.

In the first section of this chapter I discuss how prominent Strawsonians like McKenna, Vargas, and McGeer respond to the blame-sceptic by claiming that blaming sentiments are justified if their addressee deserves it, because -in those cases- these sentiments successfully communicate moral censure to the addressee and (in a way that is derived from this communication) tend to cultivate moral agency in the addressee. Within this tradition, above threshold reason-responsiveness is the paradigm agency condition for being a deserving target of blaming sentiments. From now on I refer to those accounts that justify practices of deserved blaming on the basis of their communicative and (derived) moral agency cultivating purpose, as CC accounts.

I then move on to discuss how, in their focus on blame, CC accounts leave important parts of our moral practices unaddressed. This may explain the confusing and implausible suggestions about our engagement with underdeveloped or compromised individuals. In section 2 I discuss reproach as a class of critical moral address and analyse how individuals whose reason-responsiveness is compromised or underdeveloped can be responsive to such address. In the third section I analyse why and how some compromised and underdeveloped agents can be appropriately reproached-without being deserving of blame.

I propose CC accounts include these individuals as persons who can, in some relevant, way be held responsible, and illustrate how this would resolve ambiguities in the accounts of McKenna, Vargas and McGeer respectively.

The fourth section anticipates an objection to my proposal. On certain CC accounts one could argue that appropriate reproach and deserved blame are to be considered co-extensive. I argue this response would be problematic because it facilitates misrecognition of agency, complicates philosophical debate, and undermines a theory's critical potential.

4.1 Communication and cultivation; a response to the blame-sceptic

Since Strawson wrote 'Freedom and resentment', philosophers have pondered over the normative criteria for the appropriate experience and expression of blame towards a person (P. F. Strawson, 2008). Strawsonians typically consider indignation or resentment to be the prototypical forms of blame. They ask under what conditions this blame is *deserved* by the target of these sentiments i.e. under what conditions the target is blameworthy. One of the main agreed upon criteria for being a deserving target of negative reactive attitudes is that the addressee did something that is morally objectionable.

Behaving in a morally objectionable manner is, however, not sufficient for being a deserving target of other people's blaming attitudes. There are further conditions that then need to be met for negative reactive attitudes to be a *deserved* response to a person (A. M. Smith, 2007; Coates & Tognazzini, 2013; Todd, 2012). One of the main further conditions for being a deserving target of someone's resentment or indignation is what one may call the *agency condition*. The addressed person should have certain psychological abilities in order to be a deserving target of reproachful attitudes. Strawson writes about how we are invited to repudiate and withdraw our blaming sentiments when someone's abilities to comply with moral norms and expectations are underdeveloped or compromised (P. F. Strawson, 2008, pp. 9–10).

Following Strawson, philosophers have taken up the challenge to identify the specific agential abilities that are necessary for being a deserving target of blaming sentiments. Most agree that a good account of the abilities that are required involves at least above-threshold responsiveness to moral reasons; only when a sufficiently *reason-responsive* agent fails to take into account moral considerations can this agent then be said to express the ill will or display the moral disregard that renders the person deserving of felt and expressed resentment or indignation. But if someone's reasons-responsiveness was underdeveloped or compromised, feeling and expressing blame towards the person would not be a deserved response (Fischer & Ravizza, 1998; McGeer & Pettit, 2015; M. McKenna, 2012; Vargas, 2013a; Wallace, 1996).

Reasons-responsiveness involves two types of abilities: the epistemic ability to recognize and weigh moral considerations and the volitional ability to guide one's actions in the light of these moral considerations. An agent is *sufficiently* responsive to reasons when she responds to these reasons in a range of relevantly similar circumstances. It is often unclear how large this range should be and which circumstances count as relevantly similar. I take it that the threshold range of relevantly similar circumstances is meant to single out a level of autonomy that allows us to say the person is *already able* to respond to the moral reasons that were disregarded, and of whom we may therefore expect *more* good will and regard. The person then need not go through

any further development, training or therapy in order to be able to respond to the moral reasons that were disregarded in the type of circumstance in which they were disregarded. That an agent has this ability is suggested by the agent's responsiveness to these reasons in a range of relevantly similar circumstances.

But the sceptically inclined may still ask why above threshold responsiveness provides us with the level of competence required for deserved blame. In order to respond to the sceptic, further justifications need to be given to support the claim that expressing indignation or anger to someone who *can* – in a range of relevantly similar circumstances – do otherwise, can be deserved.

One powerful response to the sceptic is that these emotional exchanges of resentment and indignation are a form of moral address that –through communication- tends to cultivate reasons-responsiveness. They communicate that the person did something morally objectionable and tend to sustain and enhance the addressee's ability to discern these moral considerations and to act in the light of them i.e. they elicit uptake (Macnamara, 2015a). On these accounts, deserving targets of blaming sentiments are said to be generally understanding of what these emotions convey and expressing these emotions is said to render the addressee more able to discern and be motivated by these considerations in the future. These observations are taken to be broadly consistent with central evolutionary and naturalist accounts of the development and function of moral anger in psychology and cognitive science (Keltner, D. & Haidt, J., 2003; Lewis et al., 2010; McGeer, 2011; Nichols, 2007; M. Vargas, 2013).

For those philosophers, a reactive attitude is an apt move in a conversation if it is directed at a person who can understand and respond to reactive attitudes as a form of moral address. Throughout this chapter I will use the word appropriate to refer to this communicative aptness. It is generally assumed that individuals who have above threshold reason-responsiveness can typically also understand and respond to blame as a form of address, which is to say that if blame is deserved it is also appropriate.

Together, the value of communication and cultivation are purported to explain to the sceptic why felt and expressed blaming attitudes -to a competent agent- can be deserved: when they track an agent who is deserving of them, they communicate and tend to cultivate. According to some philosophers appropriate address of a person who expressed ill will or moral disregard, in itself provides the required further justification for why some people deserve blame (M. McKenna, 2012; Watson, 2004). For cultivation theorists like Vargas and McGeer, the moral

agency-cultivating *outcome* of appropriate address is the focal point of justification for deserved blame (McGeer, 2014; McGeer & Pettit, 2015; Vargas, 2013a).

4.2 Attending to reproachful engagements with underdeveloped or compromised agents

In their response to the sceptic, the above accounts primarily focus on blame. These accounts thereby leave other relevant types of critical moral address under-analysed. Less charitably, other forms of critical moral address, on these accounts, are not only left unaddressed but also seem to be misrepresented. Strawsonians often seem to suggest that (sincere) critical and emotionally engaged forms of moral address are neither natural or desirable responses when an addressee lacks above threshold reason-responsiveness (Shoemaker, 2007; P. F. Strawson, 2008; Wallace, 1996; Watson, 1993). In reality the conduct of underdeveloped or compromised agents does affect us, and we do at times have critical moral conversations with them as a result of being negatively affected by their conduct. In this chapter I argue such interactions can be appropriate and discuss how we should conceive of such reproachful engagements on CC accounts.

The sort of reproachful engagements I have in mind are in important ways akin to the types of blaming that are central to CC accounts. In this section I provide an account of reproachful attitudes that is similar to, but does not completely overlap with (what are typically considered to be) blaming attitudes. I then discuss how compromised or underdeveloped agents can be sensitive to the morally communicative function of such reproachful attitudes and therefore would be appropriately reproached even though they would not deserve to be blamed.

A reproachful attitude is an affective response to someone's behaviour that exists in an appraisal of this behaviour as norm-transgressive, and a bodily readiness to in some way challenge the agent, or interfere with the agent (Keltner, D. & Haidt, J., 2003; Lewis et al., 2010; Nichols, 2007). Reproachful sentiments are thus tied to the (perceived) transgression of an interpersonal norm. They also typically manifest that we care about this norm.

Reproachful attitudes may be expressed by something as slight as a frown, a raised eyebrow or a serious stare. They may also be subtly couched in a surprised response or a question tinged with disapproval. Our expressions of reproach in response to other people's behaviour are incredibly multifarious, and people are typically quite sensitive to the variety of facial, verbal and bodily expressions of reproach. But it should be noted that reproach, as I understand it, constitutes critical address, engages the addressee, and calls for some sort of response from the addressee.

Obvious as this may seem, it is important to point this out because it distinguishes reproachful attitudes from so called ‘attributability responses’ (Shoemaker, 2015). An attributability response does not necessarily serve to engage and address the person who elicited this response. These responses only concern an attribution of fault or of certain undesirable traits to the person. Such responses *only* morally evaluate a person, but they do not call for a response from the person who is evaluated. Typical examples of attributability responses are contempt, disdain or revulsion. I set such responses aside and only focus on emotional responses to a person that are associated with critical address and engagement. These responses do more than just critically evaluate the person; they also aim to address and engage the person.

In this sense, reproachful attitudes are especially similar to blaming attitudes on CC accounts. Vargas for example writes that blame “is very much like a marking of morally salient considerations. It doesn’t work by unfettered manipulation of behaviour disconnected from any invitation to reflect on the involved moral considerations (Vargas, 2013a, p. 202).’ McKenna writes that blame, among other things, communicates our moral expectations and disapproval to the addressee (M. McKenna, 2012, Chapter 7). These observations would also apply to reproachful attitudes more broadly. They are a response to morally objectionable conduct that communicates the relevance of certain moral considerations to their addressee and invites the addressee to reflect on what is morally acceptable and what not. Just like blame on CC accounts, a reproachful attitude is an affective response which, when expressed, communicates the relevance of some moral norm and invites the addressee to reflect on her conduct (e.g. McKenna, 2012, Chapter 7; Vargas, 2013, Chapter 8).

But reproachful attitudes are not identical to blaming attitudes. Blaming attitudes are a particular subcategory of the larger class of reproachful attitudes. The message of reproach, broadly conceived, can carry a wide range of contents.⁶ The communicative message of reproach, among other things, may convey that you did something that is unfair, disrespectful, or in another way morally objectionable within a given community.⁷ It aims to convey some type of disapproval to the person in response to her not complying with some interpersonal norms, standards or expectations that you have. It is in this regard that reproach, broadly conceived, is different from (most understandings of) blaming sentiments. Reproach can convey to the person that she acted objectionably *without* expressing that the person is deserving of blame and without attributing blameworthiness to the person. Contrary to blame, reproach does not necessarily also

⁶ Note also how other feelings, like sadness or laughter, may at times convey reproach and even blame.

⁷ An angry response need not even have *moral* content. When one gets angry at a dog this probably only conveys a threat or command like: ‘comply or I’ll punish you!’ Thanks to Shaun Nichols for this suggestion.

convey that we consider the addressee's conduct to be an expression of ill will or moral disregard which requires above threshold reason-responsiveness.⁸

Some philosophers would probably object here that they do mean for the class of blaming attitudes to be co-extensive with reproachful attitudes. This seems counter-intuitive and inconsistent with most philosophical accounts of blame, but that need not be a problem if these philosophers mean to *revise* our understandings of blame. I will pay special attention to this objection at the end of the chapter. For now I will set it aside and assume that there are forms of reproach that do not qualify as blaming.

In what follows I ask if and how compromised or underdeveloped moral agents can be sensitive to the communicative and moral agency cultivating function of *these* forms of reproach that are not blame. I will argue that they can be susceptible in at least two different ways. I first discuss how their responsiveness to a moral reason may be enabled by reproachful address and then consider how their responsiveness to moral reasons can be developed by reproachful address.

The examples of interpersonal reproach one reads about in the literature often concern a response to wrongful or morally untoward actions that have already taken place. In reality we also often respond reproachfully at the time that an action is taking place. I take it that intervening reproachful responses can render the addressee responsive to a moral consideration that one would not otherwise recognize or be moved by. To illustrate, consider what happens in the following scene from a Louis Theroux documentary on people with severe brain damage. Among the persons with brain damage who are followed in this documentary is a man called Earl. At one point, Louis is talking to Earl's mum when Earl suddenly comes bursting in and is verbally abusive to her. His mum starts crying, and tells Louis this abuse happens all the time. Louis looks at Earl and asks him if he has just been playing FIFA and lost. Earl confirms that he has. Louis then asks, 'You are taking this out on your mum, aren't you?' Earl doesn't respond. Louis adds: 'I think you should apologize to your mum.' Earl asks, 'Why? She doesn't care.'

⁸ According to a number of philosophers, blaming sentiments like resentment typically involve feeling personally aggrieved, threatened in one's standing or relationship (e.g. Hieronymi, 2004; McKenna, 2012, Chapter 3; Shabo, 2012; A. Smith, 2013; Wallace, 1996).⁸ In comparison a reproachful attitude can also be more akin to frustration, annoyance or disapproval. If, for example, someone else pushes you when standing on a train platform during rush hour you may have an acute and local reproachful response to the person who pushed you (e.g. Deigh, 2011). This reproachful response may then, for example, be maintained or reinforced if you realise that the person pushed you on purpose, and turn into a feeling of resentment.

Louis then looks demonstratively at his mum (who is crying). Earl also looks at his mum, and then walks over, apologizes, and gives her a big hug (Pickup & Theroux 2016).

Earl is probably not sufficiently responsive to his mum's feelings in a range of relevantly similar circumstances. His mental capacities are severely impaired and he has been generally more abusive since his brain injury. It is likely that he fails to grasp what he is doing to people by behaving in this way when he does. This is also suggested by his remark that his mum 'doesn't care', even though she is crying right in front of him. Hence, recognizing the feelings and concerns of others as things that have bearing on how he should act doesn't seem to be something that Earl can be expected to do under circumstances like these. It is also hard for him to regulate his crankiness and to control for his impulse to take this out on others. His actions are better explained by an inability to notice and respond to certain moral considerations in these sorts of circumstances.

However, Louis responds reproachfully when he sees Earl acting in an abusive manner and gives expression to his disapproval when he addresses Earl. When we witness this interaction on the screen, we can tell that Louis is negatively affected by Earl's conduct and subtly expresses these feelings to him when calling him out. On my optimistic interpretation of this case, Earl then becomes appropriately responsive to his mum's distress because Louis has made this distress salient to him as a relevant moral consideration under the circumstances. Louis' reproach enabled Earl to respond to his mum's distress as a reason for providing her comfort and apology. Without it, Earl would not have been responsive to this moral consideration in this specific situation, nor in relevantly similar circumstances where third party reproach is absent. This reproach, in Earl's case, may also be unlikely to lead to the development of an independent ability to respond to these moral considerations in a range relevantly similar circumstances, but it enabled Earl's responsiveness to moral considerations right there and then.

One may of course refuse to go along with my optimism here. Maybe Earl is not touched by his mum's suffering at all, and only responds to Louis' reproachful sentiments because he wants Louis to like him, or feels threatened by him, or maybe his response is due to some other non-moral motivation.⁹ But even if this is what happens on this specific occasion, the possibility that someone's responsiveness to a moral consideration is at times enabled by an expression of reproachful sentiments is extremely likely. Think of all those typical situations in which someone is (temporarily or permanently) unresponsive to the harm she is doing to others, but becomes responsive to it upon being reproached.

⁹ The distinction between moral and non-moral motivations may not be as robust as suggested here: moral and non-moral motivations may overlap and be linked to one another in intricate ways (e.g. Schwitzgebel, 2017).

Some people, for example, have serious difficulty participating in meetings and presentations because they are not sensitive to the social dynamics of who may talk when and for how long. They respond to the speaker whenever an objection occurs to them and take up too much time in Q&A sessions. The sort of moral considerations we generally think one should be motivated by in these types of contexts are possibly simply not detected, rather than disregarded, by the person. The fact that other people respond reproachfully to this behaviour can, however, enable responsiveness to the entitlement of others to speak and ask questions as well.

Such enabling forms of reproach are common. In schools and clinics, cooperative activities often need to be regulated by caregivers because the participants are not always sufficiently independently responsive to the needs and concerns of other people. Some children in school will continually try to get the attention of the teacher when the teacher is trying to explain something to another child, or a patient in a psychiatric ward may tend to become verbally abusive whenever he feels powerless or fearful. These people may under these circumstances be oblivious to the impact that their behaviour has on the feelings and concerns of others. Regulating this behaviour often involves some expression of felt reproach that triggers responsiveness to the moral concerns at stake.

Reproachful sentiments are also likely to be a form of feedback that fosters the development of moral responsiveness. To see how, imagine a four-year-old and his dad travelling by train. The child is very excited and points out everything he sees to his dad. In his excitement he gets louder and louder until he is basically screaming about everything he sees, and he repeatedly pushes against his dad's shoulder to gain his attention. The dad addresses his child reproachfully: 'Shh, quiet down! Look around you. Other people are working or reading and you are bothering them!' In response to this, the child looks at his dad with his eyes wide open and steals some glances around him. For the next 15 minutes he only *whispers* excitedly about everything he sees. After that, of course, he gets louder and louder again as he is – for the first time in his life – approaching some big impressive station full of trains and people and dogs and other exciting things.

The parent in our example is affected by the child's behaviour. The child is, after all, screaming *very* loudly, which is bothersome to him and to others, and the repeated pushing is a cue for anger too. Of course these things do not *always* arouse a parent's anger, but on some occasions they do and this is one of them. The dad then expresses these feelings by pointing out to his son that he shouldn't behave that way. His anger is conveyed to the child in his facial expression and in the tone of his voice.

At the age of four, the child is probably not yet able to recognize and respond to the moral considerations at stake. He is not yet sufficiently able to independently grasp and be moved by the relevant needs and concerns of others when travelling by public transport. The child does not yet have the habits or resources for control that adults have at their disposal, especially not when his enthusiasm takes over. Hence, the screaming child does not meet the two conditions for reasons-responsiveness.

But babies and young children do have sensitivities and inclinations that are the rudiments of or – depending on one's definition – precursors of the responsiveness to moral reasons (Gopnik, 2009; Sagi & Hoffman, 1976; Smetana, 1989; Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992). Some of these seem to emerge without any relevant exposure to social feedback; infants already tend to sooth someone in pain and the differentiation between harming and helping emerges quickly (Bloom, 2012).

But social feedback is central to the further development of moral responsiveness to different types of reasons across a variety of situations. This is for example the case where it concerns sensitivities that are informed by complex cultural and technological backgrounds and contexts that complicate the instinctive grasping of the moral import of one's behaviour (like the moral importance of being quiet on a train, saying 'thank you' to the local baker, or refraining from taking sweets and shiny toys from a store). The child cannot yet independently recognize how certain considerations have moral relevance in certain types of circumstances and a parent or carer has to help them in making these connections(e.g. Nichols, Kumar, Lopez, Ayars, & Chan, 2016).

Now, of course, not all forms of social feedback concern the feeling and expression of reproachful attitudes. It is not very helpful for a child's moral development to *always* respond reproachfully to any socio-moral transgression. The literature on moral development suggests that parents should ideally employ a variety of 'disciplining strategies' that are catered to the specific child, context and type of behaviour (Grusec & Goodnow, 1994; Killen, Smetana, & Smetana, 2005).

Reproachful attitudes have a place in this variety of strategies and interactions. Research suggests that a mother's tendency to respond reproachfully in response to anti-social facts successfully fosters moral development (Grusec J.E., T Dix, R Mills, 1982). It is also suggested that the affective character of a corrective response helps to motivate a child to respond to the reasons that parents give them and helps to stress the moral importance that parents attach to certain standards (Clarke-Stewart, 1973; Grusec & Goodnow, 1994; Hoffman, M.L., 1983; Killen

et al., 2005). Children – in some sense – pick up on what reproach-expressions convey, and this understanding feeds into the development of their moral responsiveness.

Hence, though the screaming child in the train lacks above-threshold responsiveness to moral considerations under these circumstances, the child is (either in a non-moral or minimally moral sense) responsive to his dad's emotions. If parents get angry, this – at the very least – conveys to a child that something is not OK, and if they praise the child this signals that their behaviour is OK. The child's emotional responsiveness already, then, renders the expression of parents' reproachful attitudes appropriate because it is a successful form of address and cultivation. As time proceeds, the child will be increasingly able to reflectively connect these expressions to moral values and standards.

The different ways in which parents express and communicate their reproachful attitudes reflect the different developmental stages. For very young children, these communicated feelings and responses are usually purely affective and directive in character, whereas these feelings are accompanied by instructions and explanations as soon as the child is old enough to be receptive to those. When older children have developed a reasonable level of reflective awareness of moral standards, a parent's response to transgressions tends to be affectively reproachful in character *without* being accompanied by instructions and explanations (Killen et al., 2005, chapter 9). Note that the child's awareness in these situations is probably still not sufficient for above-threshold reasons-responsiveness and blameworthiness on all accounts. Nevertheless, all of these felt and expressed responses to socio-moral transgressions can be evaluated as appropriate in the light of what they convey and the moral development that they foster.

Average adults can also fall below the threshold responsiveness in a range of certain types of circumstances. There are, for example, a number of arguments in the literature in favour of responsibility-relevant impairments of volitional control over implicit attitudes (Antony, 2016; BRANDENBURG, 2016; Huebner, 2016; Levy, 2016). And people's abilities to recognize moral considerations may also be compromised in certain types of circumstances. Calhoun, for example, discusses how many people non-culpably fail to recognize that their behaviour is in some sense sexually or racially oppressive when they live in a society that is guided by norms and standards that foster this oppression. She also argues that one can nevertheless be justified in using reproach as a tool for social change in that context (Calhoun, 1989). McKenna agrees with her but notes that, though justified, these attitudes would not be *fitting* as '*moral address*' to the agent, given the agent's inability to recognize forms of oppression (M. McKenna, 2017). But this form of address may still be appropriate in so far as it communicates that their behaviour is

sexually or racially oppressive, and potentially scaffolds or develops the person's responsiveness to racial or sexual oppression even if the person is not deserving of blame, i.e. despite a below threshold ability to respond to moral considerations one may still be receptive to reproachful attitudes and therefore appropriately be addressed by them. If, say, a married woman in the 1960s angrily confronted her husband with the injustices he was unwittingly subjecting her to, this form of address may enable or develop his responsiveness to these concerns. His sensitivity to what her feelings convey to him would then render him –in the communicative sense of the word- an appropriate addressee of reproach, but he is not thereby also a deserving candidate of blame.

4.3 An addition to CC accounts

If the skills for responding to moral considerations and responding to blame are pluralistic in character and come in degrees, it is plausible to assume that many of those who fall below the threshold for reason-responsiveness may be in possession of some preliminary aspects of these skills and/or have these skills to *some* degree. This assumption finds further credence when we look at our actual interactions with underdeveloped and compromised agents, and observe what they can do.

The agents discussed in the last section clearly have some agential abilities. They are responsive to a moral message that reproach (broadly conceived) conveys. And because this reproach conveys the salience of a moral consideration, the agents are *also* responsive to a moral consideration when reproached. But it seems wrong to say of them that they are *already sufficiently able* to respond to these moral considerations by themselves. They lack the type and level of agency required for blameworthiness for their conduct. Without reproach as a form of enablement and development of this responsiveness, these agents are *hardly ever* attuned to such reasons under relevantly similar circumstances (where this does not include those circumstances in which one's responsiveness is enabled by another person's reproach). I propose CC accounts add an agency criterion to their theories that would –all other things being equal- suffice for being an appropriate addressee of forms of reproach that are not blame.

On the picture I propose, above threshold reason-responsiveness still is a necessary agency condition for being a deserving addressee of blame. But in addition to this, the minimally required agency condition for being an appropriate addressee of reproach more broadly conceived is responsiveness to reproach. This condition does not refer to abilities to abide by moral norms, but specifically refers to abilities to understand and respond to moral conversation.

An agent is already responsive to reproach when her responsiveness to a moral consideration can be enabled or developed *through* communication of reproachful attitudes. Compared with the above threshold reasons-responsiveness requirement, the condition of reproach-responsiveness is met if:

Being addressed by reproachful attitudes a) renders the moral reasons one did not respond to (more) salient to the person now or in the future and b) may motivate the person to respond to this reason now or in the future.

If communication and cultivation models are right, then above threshold reason-responsive agents are typically *also* responsive to reproach. But responsiveness to reproach in itself falls below the threshold of responsiveness to reasons required for blameworthiness. The agents discussed in section 3 are, for example, *only* responsive to reproach and not (or not yet) independently sensitive to the moral considerations that this reproach renders salient. They are appropriate addressees of reproach, but not deserving targets of blame.

I relatedly propose that there is then an important way in which reproach-responsive agents *are* responsible, and may be *held responsible* for morally objectionable behaviour. They may be held responsible in so far as they may be called out on this behaviour and may be called upon to do better now or in the future. They are responsible in so far as they are capable of minimal moral communication and sensitive to the moral agency-cultivation of reproach. This forward-looking sense in which these agents are responsible and can be held responsible, renders them in a relevant way participants in our moral practices and appropriate subjects for moral address (BRANDENBURG 2017 vs. Pickard, 2013). Below I will discuss how reproach-responsiveness clarifies and supplements the theories of respectively Michael McKenna, Victoria McGeer and Manuel Vargas.

McKenna, in agreement with Watson, says that one of the crucial abilities that an exempted agent lacks and a responsible agent has is ‘a capacity to be addressed through the morally reactive attitudes and their attendant practices.’(M. McKenna, 2012, Chapter 4) On his account an exempted agent is ill equipped to understand what is communicated through the manifestation of blaming attitudes.

But young children and cognitively compromised agents *can* often be addressed through expressions of reproachful attitudes. The reproachful expressions in the above-discussed scenarios are, for example, communicatively felicitous. And these agents may even be able to

understand (parts of) the message of the deeper forms of reproach that we characterize as resentment and indignation (despite the fact that they are not deserving of those responses).

Does this then mean that these agents are not exempted from responsibility on McKenna's account? Because they can in an important way participate in a moral conversation - they can understand and respond to other people's reproach- it seems they are not suitable candidates for McKenna's interpretation of Strawsonian exemption.

But they are not fully responsible agents on McKenna's account either because they do not meet the agency criterion for being a *deserving* target of blaming sentiments. The criterion for blameworthiness, on his account, is that a person knows her action is wrong, and performs it freely (M. McKenna, 2012, Chapter 4). But how could an agent not be exempted, but not be responsible either? McKenna's account is unclear when it comes to the question of how we are to conceive of agents that are reproach responsive but not (above threshold) reason-responsive.

By allowing for my supplement McKenna could include a class of agents for whom reproach could be a communicatively appropriate form of moral address. Reproach-responsive agents can in a relevant way participate in responsibility practices because these agents can understand and respond to reproach. He could also still maintain that these agents are not always appropriate *and* deserving addressees of the particularly *strong* forms of reproach that qualify as blame.

McKenna could explain this because the *message* of blame is different from the message of milder forms of reproach. Blame conveys –among other things- that we attribute to you a type of ill will or disregard for others that requires above threshold reason responsiveness. This would only be warranted if the person meets the required threshold. When an addressee is reproach responsive *only*, milder forms of reproach would be the warranted and communicatively felicitous manner of address.

Reproach-responsiveness is also a relevant addition to Vargas' agency-cultivation model. Vargas paints 'a picture of desert on which responsible agents deserve reactions to their blameworthy actions because such reactions help aid agents in their self-governance in light of moral considerations (Vargas, 2013a, Chapter 8).' According to Vargas, in the ordinary case, blaming 'tends to push the target's attention to considerations that others perceive as morally salient. In this way, blame connects agents with psychologies like ours to moral considerations (Vargas, 2013a, Chapter 8).'

Please note how similar this description is to the effect that reproachful attitudes tend to have on a reproach-responsive addressee. By reproaching a reproach-responsive addressee, one

also pushes the addressees attention to considerations that others perceive as morally salient and this address typically has a derived agency-cultivating function too. In this regard appropriate reproach only differs from deserved blaming in so far as it develops or enables rather than maintains or enhances the addressee's agency.

But when it comes to the ways in which we can appropriately make moral considerations salient to non-culpable agents who fall below the threshold for reason-responsiveness, Vargas is silent. He only provides a brief and slightly puzzling description of our responsibility attributions to children. He claims that children below the reasons-responsiveness threshold are typically subject to *feigned* responsibility attributions (Vargas, 2013, Chapter chapter 7, section 8). I don't think our interactions with children or compromised agents are typically 'feigned' interactions. Surely when we reproach children, this reproach is not always heartfelt. The naughty behaviour of a little child at times makes us laugh rather than feel reproachful. In such cases we may nonetheless pretend to feel reproachful. But there are also times when we really do feel reproachful towards children and express these feelings to them.

What is it then that Vargas is getting at when he says our responsibility attributions to children are generally 'feigned'? My suspicion is that Vargas would describe the reproach of a child as a *feigned* responsibility attribution because he assumes this is a form of holding someone responsible and also assumes that all forms of holding responsible are blame attributions. We rarely believe that young children are sufficiently reasons-responsive to be deserving of blame for what they did. What is in fact feigned according to him is, then, not so much the emotional expression but rather the blameworthiness attribution.

He hereby overlooks the possibility that we can *genuinely* reproach someone without also blaming the person. An attribution of blameworthiness and above-threshold reasons-responsiveness is not necessarily implied in expressing reproachful attitudes to someone. The child thereby *is* responsible in *some* of Vargas's sense of the word, because the child is sensitive to the cultivation function of reproachful attitudes. This sense of 'responsibility' renders the child an appropriate target of these attitudes. This theory would therefore benefit from including reproach-responsiveness as a type of agency that suffices for some forms of holding responsible i.e. that suffices for forms of genuine reproach that do not amount to blame.

Also on McGeer's cultivation-model it is not immediately clear how we are to think of children or compromised agents who are responsive to reproach but are not independently able to respond to moral reasons. According to McGeer, 'reactive emotions are those emotions we think it appropriate to feel only towards "responsible agents"'(McGeer, 2011). On her account these

reactive emotions encompass the class of reproachful attitudes broadly conceived. But elsewhere, McGeer and Pettit write that on the normative theory they propose, ‘the features that explain failure without counting as excuses are just those factors—those glitches and chances—that are susceptible, according to our injunctive assumptions, to the regulatory effects of our holding one another responsible’ (McGeer & Pettit, 2015, p. 29). Hence, if one is without excuse one is an appropriate addressee of blame because culpability goes hand-in-hand with a susceptibility to the capacitating function of resentment and indignation. In contrast, exempting factors imply that a person is *not* susceptible to the moral-agency-cultivating effects of the reproachful attitudes.

The problem on this account, it seems to me, is not that we cannot account for reproach-responsive agents but that we cannot account for the difference between reproach-responsive agents only and agents who meet the threshold for reason-responsiveness. A distinction between reproach-responsiveness and above threshold reasons-responsiveness can help this theory because it marks and explains the different *types* of responsibility responses that we can find in our practices. We would reproach a young child who behaves objectionably and is reproach responsive, but we would only blame a friend or colleague who can *already* respond to the reasons that she disregarded.

4.4 Anticipating an objection

The at times interchangeable use of the terms ‘blame’ and ‘moral address’ on CC accounts may suggest to the reader that the class of reproach and the class of blame are to be considered co-extensive on these accounts. I have so far assumed that this is not the case and that the concept of reproach and reproach-responsiveness would be a useful addition to these theories. For most accounts this seems plausible given their descriptions of blame as referring to specifically resentment or indignation and being accompanied by attributions of ill will or disregard for moral reasons. But certain functional accounts of responsibility may suggest that we change what we *mean* by blaming sentiments so as to include any type of moral reproach.

McGeer especially at times suggests that blame is co-extensive with reproach. On her account, reactive attitudes refer to a broad range of affective disapprovals we may experience in response to perceived norm-transgressions (McGeer, 2013, p. 171). In a later chapter she discusses how being *genuinely deserving* of blame on her account means that you are the sort of agent who is due the respect of being seen as a participant who is sensitive to the scaffolding function of reactive attitudes (McGeer, 2018). Reproach-responsive agents are due this respect

too, so it seems that on her account these agent's would be deserving of blame. Maybe other theorists would opt for a similar response to my proposed addition.¹⁰

My argument has shown that if one were to equate reproach and blame, one would also have to lower the agency criterion for being an appropriate addressee of blame. Reproach responsiveness suffices for being sensitive to the communicative and cultivating function of reproach broadly conceived. But I advise against this move for three reasons. I think the distinction between appropriate reproach and deserved blame usefully refers to different types of evaluations of an agent that can be found in our normative practices, facilitates philosophical debate about these practices, and allows for criticism of these practices. Below I elaborate on these three reasons for maintaining the distinction.

McGeer writes about the sort of capacities that are tracked by blaming attitudes:

It is certainly part of our message that we expect, and indeed, demand, that individuals show one another an appropriate degree of moral regard. But given that our reactive attitudes are sensitive to judgments that we make about whether or not someone is a fitting recipient of these attitudes, the fact that we express them effectively communicates a good deal more. It says to the recipients that we don't despair of them as moral agents; that we don't view them 'objectively' – i.e., as individuals to be managed or treated or somehow worked around; indeed, that we hold them accountable to an ideal of moral agency because we think them capable of living up to that ideal (McGeer, 2011, p. 306).

I think her diagnosis of what an expression of reproach conveys to someone is roughly correct here. When we reproach someone, we implicitly convey that we take someone to be able to live up to some sort of expectation or norm. But the specific *type* of ability to live up to our expectations that we are conveying when we reproach someone matters quite a lot. Conveying or implying that someone 'is already able to do otherwise', is decisively different from conveying that someone can come to live up to moral norms or expectations (BRANDENBURG, 2017). Even though both messages may engage and capacitate the addressee, the first serves to sustain or enhance moral agency whereas the second is a form of enabling or developing moral agency.

¹⁰ Miranda Fricker, for example, seems to subscribe to a broad notion of blame when she equates paradigmatic blame with any type of moral correction that achieves shared moral understanding through communication (Fricker, 2016). Maureen Sie may also have a broad category of blaming sentiments in mind when she suggests we do and should express resentment towards young children (Sie, 2018).

In both cases we do not ‘despair’ of someone as a moral participant, but deserved blame is traditionally only associated with the first form of address.

Minimally reproach-responsive agents would be miscategorised by such an account. The recognition that is internal to addressing these agents through reproach is recognition of the person as someone who minimally understands and is responsive to your reproachful response. This response evaluates the other as someone who cares about and is responsive to how what she did affected us. But when we take someone to be deserving of blame, we evaluate the other person as someone who is already able to live up to a moral norm but who has been disregardful of this norm. One reason to distinguish between reproach and blame is to account for these different evaluations of the person, which can accompany or be embedded in a reproachful response.

Please note that this is consistent with the observation that, in our practices, it is not always clear which one of the two one is engaging in. People may waver between reproach and full on blame, or they may be unsure about how they exactly think of the addressee. The conceptual distinction between appropriate reproach and deserved blame in fact helps us to voice such indecisiveness or doubt.

Another reason to distinguish between blame and reproach is for clarity of philosophical debate. Stressing the distinction brings into sharp relief what the disagreement between blame-sceptics and CC accounts is actually about. Blame sceptics like Pereboom or Milam claim that -on their typical definitions- blaming attitudes such as resentment or indignation should ideally be abolished because they are harsh and hostile and no person in our world has what it takes to be a *deserving* target of such attitudes (Milam, 2017; Pereboom, 2009). But the fact that no one is a deserving target of blame does not also mean that it is never in any other way appropriate to address someone by means of reproach. The sort of abilities that would suffice for mild and appropriate forms of reproach differ from the degree of freedom that the sceptics consider to be required for deserved blame (See for example Pereboom, 2014).

So, in response to the sceptic CC accounts have to argue that there is something valuable about *deserved blame* where this refers to the narrow and strong understanding of blaming attitudes, and not just any type of reproach. If a CC account does not distinguish between appropriate reproach and deserved blame it is unclear whether the narrow and strong notion of blaming attitudes is in fact defended on this account. For example, on McGeer’s account it is unclear whether moral communication and cultivation are *essential* to the stronger blaming attitudes, or may just as well be secured by milder forms of reproach. If strong blaming attitudes

are not essential for establishing moral communication and moral agency cultivation, McGeer has not in fact deflected the blame sceptic.

A third reason for distinguishing between appropriate reproach and deserved blame becomes apparent now: if one doesn't distinguish between appropriate reproach and deserved blame, one loses critical potential. On occasions it may be important to criticise someone for unduly blaming a person *without* thereby denying that this person may be appropriately reproached. If bystanders leave someone wounded in the streets, or if people continue to administer electric shocks to someone when a man in a white coat tells them to do so, expressing reproach as a witness may be an apt form of address communicating the moral import of these transgressions and facilitating moral uptake. But if it is true that these people were not *already able* to respond to the relevant moral reasons under a range of relevantly similar situations, this should give us pause when it comes feeling indignant and *blaming* them for what they did.

On other occasions it may be important to criticise someone for withdrawing and repudiating her reproachful responses towards someone based on inabilities that do not in fact justify such a withdrawal of engagement. If the person towards whom reproach is felt *would* care to know about and be responsive to these feelings, one may have reason to express these feelings despite the person's inability to independently already recognize and respond to a moral reason. The distinction between appropriate reproach and deserved blame allows us to voice these criticisms.

4.5 Conclusion

In this chapter I have proposed an addition to CC accounts. CC accounts respond to the blame-sceptic by arguing that blame is a form of address that tends to cultivate moral understanding and, as a result of that, moral agency in their addressee. By focusing on blame these accounts leave other interesting aspects of our moral practices unattended. I pointed out that reproach more broadly conceived has a very similar role. Reproachful attitudes are akin to blaming sentiments and they may successfully communicate with- and cultivate the agency of- underdeveloped or compromised agents.

I discussed how adding the criterion of reproach-responsiveness to CC accounts as an agency-criterion sufficient for reproach (more broadly conceived) usefully supplements and clarifies these accounts. It allows these theories to account for a class of agents who fall below the threshold of reason-responsiveness required for blame, but can still participate in our moral practices and conversations in relevant ways.

A revisionary CC account could object that reproachful attitudes equate to blaming attitudes. I argued this objection would be problematic because the distinction between appropriate reproach and deserved blame usefully refers to different types of moral evaluations in our practices, clarifies philosophical debate, and gives a theory critical potential.

I therefore suggest that CC accounts incorporate a class of agents who are insufficiently competent to be deserving of blame but sufficiently competent to be appropriate addressees of reproach and, in some relevant ways, participants in our responsibility practices.

5 A Case Study on Strawsonian Exemption from Responsibility

Abstract

This chapter evaluates the Strawsonian notion of exemption from responsibility by asking whether it correctly applies to the group of people diagnosed as highly functioning autistic individuals. Strawsonian exemption is generally characterized by the suggestion that incapacitated sensitivity to reactive attitudes robustly tracks or constitutes incapacitated reason-responsiveness. This connection is of theoretical interest because it implies that sensitivity to reactive attitudes informs us about a person's responsible agency. I ask what the autistic deficit in affective attunement implies for both a person's sensitivity to reactive attitudes and a person's responsiveness to moral reasons, and thereby test the purported connection between these two capacities. First, I argue a demanding definition of sensitivity to reactive attitudes would unduly exclude autistic individuals from the class of morally responsible agents. Second, I argue that when incapacitated reason-responsiveness can be local, the connection would unduly exclude autistic individuals from the set of individuals sensitive to moral address, when sensitivity is liberally defined. In order to avoid unduly excluding individuals with high functioning autism from the moral community, Strawsonians need to adhere to a liberal notion of sensitivity to reactive attitudes, and a liberal notion of reason-responsiveness. But doing so, also calls for a reconsideration of the theoretical import of a connection between these two capacities.

It is of great theoretical and practical importance to gain a better understanding of how one can and should morally relate to people who have been diagnosed with a mental disorder. Strawsonian theories of moral responsibility are front runners when it comes to engagement with this question. These theories give a central place to our ‘responsibility responses’, or so-called ‘reactive attitudes’, to one another, and ask under what conditions these responses are appropriate and valuable. Appropriate reactive attitudes are thought to robustly track or constitute what it is to be a responsible agent. But even Strawsonians have only recently begun to ask how the suspension of reactive attitudes towards a person, should apply to the range of different mental disorders that people may be diagnosed with. This chapter critically evaluates accounts of Strawsonian exemption by asking how they would apply to the case study on high functioning autism.¹¹

The first section of the chapter discusses how Strawsonian exemption is characterized by the idea that insensitivity to reactive attitudes goes hand in hand with incapacitated reason-responsiveness. I refer to the purported connection between these two capacities as the communication connection. According to this connection abilities to communicate in terms of reactive attitudes robustly track or even constitute sorts of capacities that are required for responsibility. I test this connection against the case study of autism and show that it is false on demanding readings of the respective capacities. I do not believe all Strawsonians endorse these demanding readings. Rather, I take it they may get into trouble once they are forced to explicitly avoid the exclusionary implications of these readings.

I take it that according to the communication connection the following two claims would hold:

The talk-walk claim: Insensitivity to reactive attitudes indicates incapacity to respond to moral reasons

The walk-talk claim: Incapacity to respond to moral reasons indicates insensitivity to reactive attitudes

This chapter criticizes two particular interpretations of these claims separately. The talk-walk claim will be evaluated in the two sections that follow. First, I engage with works by Stout and Shoemaker, which present a demanding definition of sensitivity to reactive attitudes and

¹¹ My claims about the abilities of people with high functioning autism generally apply to the group of people that are diagnosed as such but need not apply to every individual.

which argue that highly functioning autistic agents are not sensitive to reactive attitudes and are consequently exempted from being accountable. I argue that the claim that ‘insensitivity to reactive attitudes indicates an incapacity to respond to moral reasons’ is both implausible and unduly exclusive on this reading of sensitivity to reactive attitudes.

The first challenge is to avoid unduly denying that individuals with autism are responsive to reasons. I suggest one may avoid this challenge by providing a more liberal definition of what it means to be sensitive to reactive attitudes.

The second half of the chapter goes on to evaluate the walk-talk claim and argues that, when responsiveness to reasons is demandingly defined, the case study of autism presents a challenge to this second claim.

On a demanding interpretation both local and global incapacities to respond to reasons, count as incapacitated reason-responsiveness. Highly functioning autistic individuals may be locally incapable of responding to certain moral reasons. But despite this local incapacity to respond to moral reasons individuals with autism, and possibly other persons with mental disabilities, can be sufficiently sensitive to the attitudes that figure in responsibility conversations *about* these reasons. Especially when sensitivity to reactive attitudes is liberally defined.

The second challenge is then to avoid unduly denying that individuals with autism, or other persons with mental disabilities, are unable to participate in responsibility conversations. Meeting both challenges means the theoretical import of the connection may have to be revised. I suggest that, rather than informing us about the capacities that are required for responsibility, sensitivity to reactive attitudes indicates that there is potential for development of the capacities that are required for responsibility.

5.1 Strawsonian exemption from responsibility and the coincidence claim

When a person is exempted from having responsibility, Strawsonians claim, that we tend to – and should – adjust our reactive attitudes towards the person accordingly. Negative reactive attitudes like resentment or indignation are thought to appraise a person’s expression of ill will towards, or lack of consideration for, other people and the community. Generally, Strawsonians believe that a person cannot really display such ill will or neglect for others if the person *lacks* certain agential capacities required for displaying moral regard for others. Suppose that a person treats another person in a manner that is threatening or offensive but it turns out that this person is at that time actively psychotic and is therefore not able to grasp that he is treating another human being in a way that constitutes undue suffering or harm, let alone control his behaviour. This person at that point in time lacks the relevant agential capacities required for displaying

moral regard for others and therefore cannot be said to express ill will or display disregard in acting as he does.

Reactive attitudes such as resentment or indignation would in this case incorrectly appraise this person as displaying disregard, just as being scared by a coat stand that one mistook to be a strange man in the dark would incorrectly appraise the coat stand as presenting a danger. To put it simply, if a person's moral agency is underdeveloped or compromised, the correctness conditions for feeling reactive attitudes towards this person are not met (e.g. D'Arms & Jacobson, 2000). The first feature of Strawsonian accounts of exemption from responsibility is that the appraisal central to reactive attitudes (i.e. disregard for others or ill will towards others) would be incorrect. I refer to this feature as incorrect appraisal.

However, in the case of Strawsonian exemption, there is also, typically, another way in which reactive attitudes are taken to be inappropriate responses to the exempted person. According to a respectable Strawsonian tradition, an exemption from having responsibility is connected to the inappropriateness of reactive attitudes as a form of engagement and address. This second feature, the inaptness of moral address, requires more explanation.

Strawson himself suggests that it would be civilized and natural to have an 'objective stance' towards those people who are not able to meet moral expectations for mutual regard. This stance excludes 'the range of reactive feelings and attitudes, which belong to involvement, or participation with others in inter-personal human relationships' (P. F. Strawson, 2008, p. 10). He adds that from this stance the other is not seen as someone whom one can reason with or quarrel with. One can at most *pretend* to do so.

In line with Strawson's own position, a respectable Strawsonian tradition also suggests that our stance towards the exempted person involves more than an adjustment of our incorrect appraisal of her. Philosophers writing in this tradition claim that individuals who are exempted are, among other things, unfit for interpersonal involvement, incapable of having adult relationships, and incapable of moral communication and participation (e.g. M. McKenna, 2012; Shoemaker, 2011, 2015; Wallace, 1996; Watson, 1993). Shoemaker summarizes Strawsonians like Watson, Stern, Darwall, and many others to defend the following position:

'If one lacks the capacity for this sort of reason-based address and exchange—a specific type of normative competence—then one must be outside the boundaries of the moral community, beyond the reach of a certain type of relationship. And while having this capacity may not be sufficient for rendering actual praise and blame appropriate (there may be other excuses that cause us to suspend such reactions in particular cases), it is nevertheless both

necessary and sufficient for being a member of the moral community, rendering one at least eligible for praise and blame (Shoemaker, 2007, p. 71).’

And Paul Russell, for example, writes, “The responsible agent must be able to feel and understand moral sentiments or reactive attitudes (Russell, 2004, p. 293).’

This brings us to a second distinguishing feature of Strawsonian exemption: the inaptness of reactive attitudes as a form of moral address. Reactive attitudes are an inapt form of addressing the exempted person given this person’s inability to understand and respond to reactive attitudes as a form of moral address. The idea seems to be that an exemption from having moral responsibility is in some way related to this incapacity to participate in responsibility practices or, in other words, incapacity to participate in a ‘responsibility conversation’ in which reactive attitudes are understood as moral address. I will from now on refer to this supposed relation as the communication connection.

That there is such a relation is suggested in the writings of Fisher and Ravizza, Paul Russel, Bennett Helm, David Shoemaker, and Gary Watson, among others.¹² Michael McKenna explicates the relation in his conversational theory of responsibility, and others, such as Coleen McNamara, have provided expansions of this conversational theory (M. McKenna, 2012; Macnamara, 2015b). Many more Strawsonians allude to it when they suggest that exempted people *cannot* be ‘involved’ or ‘normally related to’ in terms of reactive attitudes, and there is, to my knowledge, no Strawsonian who explicitly denies that there is a relation between the two.

The connection between sensitivity to reactive attitudes and responsiveness to reasons is theoretically interesting because it provides us with epistemic access to the actual capacities that are considered to be required for responsibility. Sensitivity to reactive attitudes and the related communicative skills can be observed in people, and responsibility conversations are part and parcel of our practices. The idea is that a person’s ability to communicate in this sense indicates that she has the capacities required for responsibility and that this sensitivity provides us with epistemic access to what capacities render a person a responsible agent. Importantly, if this is how we gain access to what these capacities are, they are metaphysically possible within our life-world; they are part of the facts ‘as we know them’ (P. F. Strawson, 2008).

Note that for this connection to be interesting in this sense the class of reason-responsive agents and the class of agents that are sensitive to reactive attitudes should (at least

¹² (Fischer & Ravizza, 2000, pp. 211–214; Helm, 2012, pp. 219 & 231; Russell, 2004, pp. 292–293; Shoemaker, 2007, pp. 70–73; Watson, 1993, pp. 126–127)

roughly) coincide. Therefore the communication connection involves at least the following two claims. First, an inability to interact and communicate with a person in terms of reactive attitudes (robustly) indicates that this person's responsiveness to moral reasons is impaired. The idea is that an inability to engage in 'moral talk' implies impairment of the ability to do the 'moral walk'. I refer to this claim as:

The talk-walk claim: Insensitivity to reactive attitudes indicates incapacity to respond to moral reasons

Secondly, the relationship is understood to (robustly) hold in the other direction too. I refer to this second claim as:

The walk-talk claim: Incapacity to respond to moral reasons indicates insensitivity to reactive attitudes

But how, exactly, are we to understand the notions of insensitivity to reactive attitudes, and incapacity to respond to reasons that figure in this connection? Given the vast amount of Strawsonian literature, it would take me too far afield to answer this question for all Strawsonian accounts. But I worry that these notions are, in general, insufficiently specified. Therefore, the communication connection can easily be, and on some accounts is, interpreted in ways that would yield problematic conclusions for the case of high functioning autism and possibly for other mental disabilities.

In this chapter I show what is at stake in defining these capacities. Two particular interpretations of these two claims have problematic and unduly exclusionary implications for people with high functioning autism. In order to avoid these implications both sensitivity to reactive attitudes and responsiveness to reasons should be liberally defined. But when one does so it becomes unclear what the theoretical import of the communication connection is.

In the following two sections I test a particular interpretation of the talk-walk claim against the case study of autism, and show why sensitivity to reactive attitudes should be liberally understood. After that I evaluate the second claim and show why responsiveness to reasons should be liberally defined too.

5.2 Testing the talk-walk claim

It is generally acknowledged that autistic individuals have trouble picking up on the particular emotional state of another person. Autistic individuals, probably because of this, will also not be affected by other people's emotional states in a typical way and will find it hard to respond to these states in a manner that these states are typically thought to call for (Baron-Cohen, 1997; Bird & Viding, 2014; Klin & Jones, 2006; Zahavi & Parnas, 2003). I will refer to this inability to implicitly grasp and be affected by other people's emotional perspective as a deficit in affective attunement. Shoemaker and Stout have recently begun to analyse how this autism-related deficit impacts a person's sensitivity to reactive attitudes.

Shoemaker argues that autistic individuals cannot fully comprehend emotional address and therefore cannot respond to reactive attitudes in a way that would be required for them to be morally accountable for their actions. To be accountable is to have the agential capacities that allow one to display due moral regard for others' feelings and interests (Shoemaker, 2015, p. 93). According to Shoemaker, this accountability is related to a capacity to understand anger as a form of moral address. He writes that:

‘...fundamental to agential anger is the action tendency to demand acknowledgment from those who we perceive to have (in)sufficiently regarded us. ... the most immediate aim is to get the offending party, when returning to his own perspective, to have his shared emotions with us transform into guilt for his transgressions (Shoemaker, 2015, p. 171).’

Because those with high functioning autism cannot implicitly grasp and share the emotional perspective of another person who is in a state of moral anger, ‘it would not be appropriate to hold them accountable and consequently they would not be accountable’ (Shoemaker, 2015, p. 171). Shoemaker claims that holding the autistic person accountable is inappropriate because ‘holding someone accountable’ amounts to having the emotional response towards a person that neurotypicals tend to have in response to moral disregard, i.e. a reactive attitude. According to Shoemaker, this response cannot be comprehended by highly functioning autistic individuals, nor will these individuals be appropriately affected by such responses. In this sense, they *cannot* be held accountable. And – given the talk-walk claim – this implies that the autistic person *is not* accountable.

Stout's argument discusses in greater detail how highly functioning autistic individuals are, according to him, unable to comprehend and care about the reactive attitudes of others. He suggests that autistic individuals, especially those with alexithymia, fail to implicitly recognize and

be appropriately affected by the reactive attitudes that come with holding someone responsible for their conduct. Alexithymia refers to an inability to identify and describe emotions; this condition often, but not always, accompanies autism (Bermond et al., 2007). Stout adds that autistic people who are hypersensitive will also have difficulties in recognizing other people's reactive attitudes as being the sort of speech act that they are (Stout, 2016).

Stout then asks how these autistic individuals would fare under McKenna's account of conversational responsibility. McKenna claims that responsible agency requires an ability to engage in a moral exchange, which means that one can recognize reactive attitudes as a speech act and can respond to them by offering pleas, justifications, or apologies. According to Stout, McKenna's requirement of being able to engage in such moral conversations should be made more demanding because, according to Stout, the alternative ways in which highly functioning autistic individuals can come to recognize and respond to the reactive attitudes of others do not make for *real* comprehension of these attitudes.

According to Stout, a highly functioning autistic person may be 'functional enough to take part in the choreographed moves, and carrying on moral conversations is an important part of this. However, the ability to function in these conversations fails to explain important factors outside of this functionality' (Stout, 2016, p. 1026). What is missing is an immediate emotional responsiveness to the feelings of others that prompt these exchanges. The autistic person does not feel '*what it is like*' to experience the reactive attitudes when she 'goes through the moves' of providing desirable responses to external markers of such attitudes. Stout concludes that a communicative theory of responsibility should not therefore mistakenly include highly functioning autistic individuals as fully responsible members of the community by overlooking their deficits in prototypical sensitivity to emotional address (Stout, 2016, p. 1026).

Both Shoemaker and Stout have a demanding definition of sensitivity to reactive attitudes. If a person cannot affectively attune to reactive attitudes, that person does not qualify as sufficiently sensitive to reactive attitudes on their accounts. They also believe that this means that highly functioning autistic individuals are not accountable members of the moral community. This means they endorse a particular reading of the talk-walk claim: insensitivity to neurotypically expressed reactive attitudes implies that reason-responsiveness is impaired in a way that provides grounds for exemption from accountability, i.e. exemption from having the moral responsibility to display due regard for the interests of others. Is the talk-walk claim, on this demanding reading, a plausible claim?

One obvious reason for doubting the plausibility of this reading of the talk-walk claim is the simple fact that highly functioning autistic individuals often *do* abide by moral norms.

Kennett has pointed out that highly functioning autistic people, unlike psychopaths, can resist their impulses and desires in the light of rules of conduct that are in the interest of others and the community. A deficit in affective attunement, according to her, only complicates discerning what those interests are; it does not amount to an inability to understand and be moved by other people's interests as worthy of consideration and respect (Kennett, 2002).

If this objection succeeds, one may hypothesize that Shoemaker and Stout, and possibly other Strawsonians, are confusing the correctness of appraisal with the aptness of address. How we aptly *address* a person may depend on this person's ability to comprehend and be affected by our (expressions of) reactive attitudes. But the aptness of this address in terms of reactive attitudes is not to be confused with the aptness of the appraisal of moral disregard that is central to these same reactive attitudes. A person who is able to abide by a moral norm would be correctly appraised as lacking due regard for this norm when she (without any good excuse) does not abide by this moral norm, *even if* this person would not be able to comprehend and respond to this same appraisal as a form of moral address. The aptness of address and the aptness of appraisal come apart in the case of high functioning autism.

Imagine, for example, that Albert, who is autistic, agrees to meet Polly for an early morning walk but sleeps in without any good excuse for doing so. This is probably rather disregarding of him. Albert and most highly functioning autistic people have learned that one shouldn't be late for appointments, but he did not pay due regard to this norm. However, if Polly feels resentful and looks at Albert reproachfully when he arrives 20 minutes late, he is unlikely to pick up on this expression *as* a speech act and tune in to what it is like for her to be in this emotional state.

Stout and Shoemaker suggest that Albert's inability to recognize and share Polly's reproach implies that Albert is not morally responsible for his conduct. But Albert may still be correctly *appraised* as having been disregarding of a moral norm. After all, he is able to grasp and abide by norms regarding punctuality and he had no excuse for being late. His insensitivity to bodily and facial expressions that reflect reactive attitudes only renders it inapt to converse with him by means of such expressions, but the inaptness of such address does not also imply the incorrectness of appraising Albert to be morally disregarding.

However, Shoemaker has a response to this objection. He suspects that Kennett's examples of highly functioning autistic individuals who are abiding by moral norms are not explained by genuine sensitivity to moral considerations. He thinks it more likely that the seemingly moral conduct of highly functioning autistic individuals is explained by their passion for order and their need to abide by the rules they have been taught, rather than by a desire to do

the right thing and to take others' interests as reason giving (Shoemaker, 2015, p. 170). That is to say, they do not abide by norms *qua* moral norms, i.e. for moral reasons.

In support of this claim, Shoemaker asks us to adopt the perspective of an autistic individual, and writes,

'If you found yourself in a world in which you could not read off from people's faces what they were feeling or the sorts of signs for interaction that others do without effort, it would likely be terrifying, especially if these other people increased the noise level in your head by yelling at you, or if they punished you for failing to conform. If so, you would surely grasp at whatever rules for such interaction that you could find and obey them rigidly. Change would be your enemy; rigid rules would give you comfort. Even if you had no idea what the grounding of those rules consisted in, or whether they counted as 'moral,' it would not matter (Shoemaker, 2015, p. 170).'

According to Shoemaker, sensitivity to the reactive attitudes of others is *required* for developing the ability to abide by norms *qua* moral norms rather than for merely prudential reasons. He would suggest that Albert from our example *may* be able to understand that he shouldn't be late because people will yell at him or do something similarly unpleasant if he arrives late; but he does not also understand *why* people will make all this noise and *how* it is disrespectful of other people's interests to be tardy, i.e. he does not grasp reactive attitudes as *moral* address.

But is implicit affective attunement to others really a *necessary* resource for the development of an ability to respond to *moral* reasons? Can one not come to duly recognize and respect other people's interests via other means? Shoemaker admits that he has no conclusive evidence for his claim, but insists that it is dubitable that one can come to recognize and abide by moral norms for moral reasons in any other way (Shoemaker, 2015, p. 169).

It seems implausible to claim, though, that affective attunement to other people is the only way in which one can gain access to moral reasons. Shoemaker cites Victoria McGeer when he discusses how a deficit in affective attunement compromises responsiveness to moral reasons. But Victoria McGeer discusses how highly functioning autistic individuals can compensate for this deficit. Highly functioning autistic people can come to genuinely register and care about the interests of other people, but they do so via an affective profile that is *different* from the affective profile that underpins (a large part of) moral agency in neurotypicals (McGeer, 2008, p. 242; See also Richman & Bidshahri, 2017)

This would explain the sort of genuine moral agency we see autistic individuals like Temple Grandin or Sean Barron engage in. Two of the rules Temple Grandin considers to be central to friendship are ‘Friends are genuinely concerned about each other’s feelings and thoughts’ and ‘Friends help each other out in times of need’ (Grandin, Barron, & Barron, 2005, p. 239). Despite the fact that Grandin has trouble attuning to subtle expressions of a friend’s feelings and thoughts, she here declares that she *cares* about a friend’s thoughts and feelings and will help friends out in times of need. The deficit in affective attunement only implies, then, that she may not always be able to pick up on *how* a moral reason manifests itself in certain circumstances, e.g. she may fail to affectively attune to a friend’s distress. It does not mean that she cannot grasp how a friend’s distress provides her with moral reason for action. When Grandin *via other means* comes to realize that a friend is distressed, this provides her with reason to be genuinely concerned and to offer help. Affective attunement sometimes, then, compromises *access* to an instantiation of moral reasons. This does not also mean that a person cannot grasp and be moved by this moral reason at all.

Something similar can be said about Sean Barron when he discusses how he had a hard time registering what was an act of sincere kindness towards him and what was not, and how he learned to cope with these difficulties. Throughout this discussion, he appears to grasp the concept of kindness and appears to care about mutual kindness. He has, for example, found insincere acts of kindness towards him hurtful, and has ‘felt wonderful’ when responding to a request for help from a friend in need (Grandin et al., 2005, p. 264).

These and other self-reports from highly functioning autistic individuals provide evidence of them being able to display due regard for the interests of others despite a deficit in affective attunement. Please also note that even neurotypical people do not always recognize and respect other people’s interests on the basis of their affective attunement to these people. For example, paying one’s taxes or donating money to charity need not be based on an affective attunement to the lived experience of the poor and destitute. It may instead be based on an acknowledgement that everyone ought to pay their fair share in contributing to public goods or ought to make a commitment to distributive justice (e.g. Singer, 1972). It would be quite a conceptual stretch to deny that one is not taking the interests of others to heart in these cases, i.e. to deny that one’s choices are based on moral considerations.

In addition to this, Richman and Bidshahri have pointed out how Shoemaker’s reading of proposition 1 is unduly exclusionary of atypical agents (Richman & Bidshahri, 2017). They write that ‘this limits the range of neurodiversity claims we can accept. We can accept lots of social, cognitive, and behavioural differences as neutral indications of diversity, but we cannot

accept differences in how people respond to transgressive actions as equally valid' (Richman & Bidshahri, 2017, p. 6). Providing accounts that can allow for diverse affective responses to moral disregard as valid responsiveness to *moral* disregard would be more inclusive of marginal agents and would thereby do more justice to them and to the neurodiversity one finds in human populations.

To conclude, on Shoemaker and Stout's demanding understanding of what it means to be sensitive to reactive attitudes, the talk-walk connection is both implausible and unduly exclusive. It is implausible because the sort of conduct that is typically thought to be respectful of other people's interests, i.e. thought to be based on *moral* considerations, would no longer count as genuinely responsive to moral reasons. It is unduly exclusive because it does not allow for diversity in affective responses to moral disregard.

5.3 The first challenge: allowing for diversity in moral communication

In the last section I discussed how the case study of autism shows us that inability to understand particular forms of address, or manners of address should not be taken to indicate a person's reason-responsiveness is impaired. If one would nevertheless want to maintain that there is a connection between sensitivity to reactive attitudes and responsiveness to reasons one should: *define sensitivity to reactive attitudes in a sufficiently liberal way, so as to avoid unduly excluding individuals who communicate in atypical ways.*

Here is how this may be done. First, one may object to Stout and Shoemaker's view that comprehending another person's emotion necessarily requires a non-inferential, direct and emphatic picking up of this emotion. Comprehension of someone's mental state may simply amount to the ability to explain and predict this mental state (Nichols & Stich, 2003; Ravenscroft, 2016). There are other types of cognitive procedures that can, on many occasions, allow one to grasp the mental states of other people in this sense. Association, inference, and stereotypes are a few examples of alternative tools that can allow one to explain and predict that another person is in a particular emotional state in a particular context (Andrews, 2012; Fiebich, 2017; Zahavi & Parnas, 2003). It is worth pointing out here that a number of philosophers would argue that even neurotypicals make as much use of these tools for picking up on reasons, as they make use of affective attunement to others (Fiebich & Coltheart, 2015).

Additionally, I propose that one should not be too restrictive when it comes to the particular *type* of felt and expressed emotional response to moral disregard that can qualify as a reactive attitude. People can be differently (negatively) affected by perceived moral disregard and

we can, in principle, allow for a variety of ways of expressing such negative affective states. All these responses qualify as reactive attitudes. Therefore, being able to grasp another person's reactive attitude may minimally amount to grasping that another person is in *some way* negatively affected by perceived moral disregard rather than having a direct and emphatic resonance with a particular emotional state. A highly functioning autistic person would therefore already be sensitive to reactive attitudes if the person can grasp that other people are in *some way* negatively affected by perceived moral disregard when this is in *some way* communicated to them.

Highly functioning autistic people are sensitive to reactive attitudes as moral address on this account because when others communicate to them that they perceive a moral rule to have been disregarded, highly functioning autistic individuals will typically understand from this that these others have been in some way negatively affected by this perceived moral disregard. This is something they can respond to by offering pleas, justifications, or apologies. Highly functioning autistic people may not be able to grasp all the emotional *nuances* of the responsibility conversations that can be found in our practices. But on a liberal account they are sufficiently sensitive to reactive attitudes to be considered able to participate in a responsibility conversation.

To illustrate, if Polly were to explain to autistic Albert that she does not appreciate him being late for their appointment, Albert would, *ceteris paribus*, be able to understand that Polly is negatively affected by perceived disregard, and he can connect this appraisal to the *moral* considerations that speak in favour of keeping promises and being on time for appointments. In response to Polly, he can do more than 'just' prudentially avoid this behaviour in the future so as to make sure Polly won't yell at him. He can *understand* that she appraises what he did to be disregarding of a moral norm and can negatively appraise his own behaviour in agreement with her because he can care about the relevant moral reasons in the way that McGeer suggests (McGeer, 2008). None of these skills regarding moral interaction and conversation are excluded by Albert's inability to affectively attune to Polly.¹³

To summarize, on a liberal definition of sensitivity to reactive attitudes, we can allow for variety in how one feels and expresses reactive attitudes, and we can allow for variety in the ways in which one can comprehend reactive attitudes. This renders a liberal reading of the talk-walk claim the more plausible reading, and, for all we know, it may very well hold true across the diversity of the human population. On this reading, rather than being sensitive to felt and expressed resentment, sensitivity requires that a person can be in *some way* negatively affected by perceived moral disregard and can understand that others are in *some way* negatively affected by perceived moral disregard.

¹³ This reading would provide McKenna with an objection to Stout's criticism.

5.4 Testing the walk-talk claim

In the sections that follow I will discuss the second claim involved in the communication connection. Recall that according to the walk-talk claim: Incapacity to respond to moral reasons (robustly) indicates insensitivity to reactive attitudes. I will argue that, once we are committed to a liberal understanding of sensitivity to reactive attitudes, a particular reading of this second claim is falsified by the case study of autism. On this reading responsiveness to moral reasons is demandingly defined: both local and global inability to respond to moral reasons count as an impairment of reason-responsiveness.

I contend that the reason-responsiveness of highly functioning autistic individuals, may be in a local sense impaired. Although these individuals can and often do abide by a broad range of moral norms, a deficit in affective attunement nevertheless compromises their ability to respond to moral reasons in *certain types* of circumstances. Kennett states that it is more difficult for autistic individuals to discern what the interests of other people are (Kennett, 2002, p. 354). And although highly functioning autistic people are incredibly resourceful when it comes to developing skills to compensate for these difficulties, we should pay close attention to the demandingness of this compensation and the extent to which it is possible.

There are a number of promising treatment strategies that can help young autistic individuals develop skills to compensate for a deficit in affective attunement to others when it comes to abiding by socio-moral expectations (White, Keonig, & Scahill, 2007). These strategies involve formulating clear and concrete social rules and teaching simple social scripts and social response scripts to enable people to recognize and abide by these rules. These strategies seem to pay off when taught over time and in a range of different contexts. But these strategies fall short of enabling highly functioning autistic agents to abide by all the moral norms that neurotypical people tend to abide by. Imagine both a neurotypical person and a highly functioning autistic person find themselves in a similar type of circumstance X, where moral norm Y applies. Compensatory strategies may fail to enable the highly functioning autistic person to abide by this norm for several reasons.

First, the effort that is involved in employing compensatory means to access other people's experiences is much higher than the effort needed by neurotypicals for affective attunement to someone else's feelings and perspectives (Bird & Viding, 2014; Zahavi & Parnas, 2003). Highly functioning autistic individuals will not be able to rely on these compensatory skills in situations where a quick and effortless response is required.

Second, compensating for affective attunement requires the formulation of many socio-moral rules and scripts. Full compensation would require teaching and learning too many socio-moral rules. The norm that applies in this particular type of circumstance may have required one rule too many.

Third, some general socio-moral expectations are too dependent on contexts and persons to be accurately translated into a clear and concise rule. If so, the person cannot rely on the strategies mentioned above.

It should also be considered that not every person will have the (same level of) opportunity to benefit from these strategies early in life. In addition, consideration should be given to the fact that an autistic person will take longer to acquire compensatory skills and may therefore still be engaged in these learning processes when neurotypicals are already able to abide by certain norms.¹⁴

For all these reasons there will be occasions when a highly functioning autistic individual cannot pick up on a moral reason because she cannot (yet) under these circumstances pick up on this reason through alternative learning or cannot (yet) abide by it by means of a social response script, through no fault of her own. On such an occasion we would incorrectly appraise the person as disregarding of moral considerations due to incapacity to respond to these moral considerations.

To illustrate this with an example, imagine an occasion on which Amelia, a highly functioning autistic individual, tells an acquaintance called Bart during a family dinner that he is too fat to find a girlfriend. She does so in response to Bart's story, which has visibly upset him, about a failed date. A neurotypical person would plausibly be thought of as being inconsiderate or even a bit malicious when making such a remark. But these terms would be inapt when describing Amelia's comment. She does not recognize the considerations that render such a blunt remark disregarding. No one has ever told her that one shouldn't remark on a person's body type unless asked, because it may hurt that person's feelings. She is thereby convinced that her blunt remark is true because she has heard someone else say this and did not pick up on the humour (though arguably in bad taste) that was underlying it; she thought the remark was an objective observation. She also fails to implicitly grasp the emotional state Bart is already in and cannot track how her remark is making Bart feel. She thus cannot (yet) recognize how these factors provide moral reason for her not to remark on a person's body weight in such a manner

¹⁴ And to the extent that these abilities can yet be developed, a nurturing stance towards this person would be more apt than an objective stance (Brandenburg, 2017).

under these circumstances. We therefore cannot expect that she will not make such a remark, nor correctly appraise her to be lacking in moral regard when she does make such a remark.

Note also that even after Amelia has learned that calling someone fat may be hurtful and that one should generally refrain from doing so, it will remain hard for her to grasp how and when there can be exceptions to this rule. If she hears friends endearingly call each other fat she may interpret this as incredibly rude. Again, this may give rise to morally complex situations in which one might easily inaptly appraise her to be meddlesome or overbearing when she corrects these friends for insulting each-other. A more detailed understanding of the person's deficit allows one to see how such conduct is not an expression of disregard for some moral norm and is – in this case – even an expression of regard for what Amelia *perceives* to be of moral concern.

So the autistic deficit in affective attunement to other people's feelings and perspectives locally impairs a person's responsiveness to moral reasons.¹⁵ This does not include those occasions when it is harder but possible for a person to rely on alternative strategies. In such a case, affective attunement makes it easier to abide by moral norms, but it is not required. The person is only incapable of responding to moral reasons in those circumstances where we cannot expect the person to rely on alternative strategies (yet) at all. We are now in a position to ask whether this impairment of reason-responsiveness indicates insensitivity to reactive attitudes; in other words, whether the walk-talk claim holds when it is tested against the case study on high functioning autism.

But before moving on to test the walk-talk claim, one might object here that the question of someone's sensitivity to other people's reactive attitudes should not arise when a person was unable to respond to the moral reasons that these reactive attitudes appraise the person to be disregarding of. Surely, if someone is unable to respond to a moral reason, the person typically does not express disregard for this moral norm when she fails to abide by it. This means that communicating to the person that you appraise her to have been regardful or disregarding of moral norms would typically communicate an *inapt* appraisal. Even if the person can understand

¹⁵ Please note that this lack of affective attunement may sometimes facilitate moral agency. In neurotypical people a visceral emotional response of another person (e.g. extreme suffering, anger or fear) may prevent someone from responding to the relevant moral considerations. This failure to be sensitive to them is not always something that the neurotypical person can compensate for. When she can't, we would readily exculpate her from blame. Similarly, we should be ready to exculpate the autistic person from failures of moral agency that are explained by their deficit in affective attunement and are impossible or very hard to compensate for. My position is compatible with neurodiversity (e.g. Jaarsma & Welin, 2012). As we shall see later in the paper, I take it that some exemptions are a part of our everyday practices and internal to a moral community of competent adults.

such an appraisal, it seems pointless and even unkind to communicate it. But this objection would be wrongheaded because it confuses the aptness of appraisal with the aptness of address.

Strawson himself understood that a communicative exchange prompted by an inapt appraisal might be an appropriate moral exchange. Sometimes we will by mistake inaptly appraise a person as displaying moral disregard and she will respond to this appraisal by pleading innocence and/or providing a legitimate excuse or justification for her conduct. Strawson writes that:

‘The offering of such pleas by the agent and their acceptance by the sufferer is something in no way opposed to, or outside the context of, ordinary inter-personal relationships and the manifestation of ordinary reactive attitudes. Since things go wrong and situations are complicated, it is an essential and integral element in the transactions which are the life of these relationships’ (P. F. Strawson, 2008, p. 8).

Responding to inapt appraisals of regard or disregard is an *integral part* of normal engagement given the *complexity* of situations and the difficulties involved in appraising the moral import of others’ actions in those situations. Often we cannot be completely certain about the aptness of our emotional appraisals. But if one has good reason to *suspect* that a person was disregarding or regarding, there is a pro tanto reason to open up a conversation in which one shares this appraisal and calls upon the other to answer for herself in the form of apologies, justifications or pleas – as laid out in McKenna’s account of moral exchange (M. McKenna, 2012 chapter 5). Such communication of (possibly inapt) appraisal is an apt form of address as long as the addressee has the capacities to participate in a responsibility conversation.

So, can the communication of inapt appraisal be appropriate when the addressee was locally incapable of responding to the relevant moral reasons? On a strong interpretation of the walk-talk claim, the answer should be no. The answer should be no because locally impaired reason responsiveness indicates that the person is insensitive to such address. But the case of autism falsifies this claim. A person with high functioning autism can understand such address (in a liberal sense of the word) even when the person was incapable of responding to the moral reasons that the addressed (maybe inaccurately) believed she had disregarded.

To illustrate, in the TV detective series *Bron/Broen*, one of the main characters, the autistic Saga Norén, often delivers bad news to family members after a murder has been discovered. She does so in a rather heartless manner: she abruptly announces the victim’s death

and gives the addressee no time to gather themselves before starting a thorough interrogation. She also gets really impatient when the person she is talking to does not immediately answer, and she does not hide her impatience. Saga is probably not, and cannot be, aware of the moral considerations that she fails to respond to. She has insufficient access to *why* her conduct is objectionable and is not capable of the finely attuned responses to others that are called for in these conversations with family members.

But despite her current inability to abide by the expectation to treat these family members with more regard for their feelings, she seems able to participate in a responsibility conversation. Her boss at some point confronts her with complaints about her conduct that victims' family members have been sending to the police department. We consequently see a McKenna-like type of moral exchange: Saga has behaved in a manner that triggers certain (inapt) reactive responses; these negative appraisals of her conduct are communicated to her; and she consequently responds by conveying to her interlocutor that she *cannot* recognize how her conduct was offensive.

Saga tells her boss that she cannot start to interact in the way that is considered desirable by going on the social communication courses her boss suggests she participate in. She is aware of her own limitations (and talents) and can reason with other people about them. And it is because of this same awareness that she takes on board her boss's concerns and decides to leave the bad-news conversations to her colleague. This response is quite plausibly based on her ability to comprehend how certain negative attitudes were central to the conversation she had with her boss.

After talking to her boss, she understands that she may inadvertently end up hurting people because she is not sufficiently sensitive to their feelings and perspectives in such situations, and she responsibly takes appropriate measures to avoid doing so. She still does not understand exactly *how* her conduct was hurtful, nor can she avoid giving offence by becoming more affectively attuned to the victim's feelings. But through this exchange she probably understands that her conduct is hurtful to others in some way, although this is inexplicable to her, and she responds to this by trying to avoid those types of situations in which she is not sensitive to the moral reasons at stake.

Saga, and other highly functioning autistic individuals, can meaningfully converse with others when these others morally address them i.e. when these others communicate to them in some way that they have been negatively affected by perceived moral disregard. They can do *even if* they have significant local incapacity to (independently) recognize the moral reasons that were disregarded.

According to Strawson, excusing pleas such as ‘I didn’t know’ or ‘I didn’t mean to’ are ‘an essential and integral element in the transactions which are the life of interpersonal relationships’ (P. F. Strawson, 2008, p. 8). Exemptions, on the other hand, are thought to exclude one from such transactions and from normal interpersonal involvement more generally. But, as we saw in the case of Saga, (local) exempting pleas such as ‘I *cannot* know’ or ‘I *cannot* mean to’ can also be pleas that are an integral element of typical transactions and exchanges among grown-ups, allowing them to organize their relationships and community so as to be duly regardful of each other’s interests.

I therefore conclude that a strong interpretation of the walk-talk claim is falsified by the case study of autism. Not *every* type of incapacity to respond to moral reasons indicates insensitivity to reactive attitudes liberally understood. Local incapacities can be isolated from a person’s sensitivity to moral address and ability to participate in responsibility conversations.

5.5 The second challenge: avoiding undue exclusion from moral conversation

Recall that on Strawsonian accounts the appropriateness of reactive attitudes is believed to either constitute or robustly track a person’s responsiveness to reasons. On a constitutive account like Shoemaker’s, this connection is interesting because a person’s responsiveness to reactive attitudes constitutes the particular capacities that are required for being responsible (Shoemaker, 2017). On tracking accounts the communication connection is interesting because the appropriateness of reactive attitudes roughly tracks the capacities that are required for responsibility (McGeer, 2018). On both types of Strawsonian accounts the class of responsible agents and the class of agents that are sensitive to moral address should roughly overlap.

But the discussion above calls for a reconsideration of the relationship between incapacity based defences and sensitivity to reactive attitudes. Local incapacities to respond to moral reasons are distinct from a person’s sensitivity to reactive attitudes, especially when such sensitivity is liberally defined. The challenge is then to define the connection in such a way that one does not unduly imply that locally incapacitated individuals cannot respond to moral address and in any way participate in responsibility conversations.

Some people’s (temporary) inability to respond to moral reasons does indeed seem to go hand in hand with insensitivity to reactive attitudes. In those cases it is inapt to communicate to them how we (incorrectly) appraise them to be disregarding of moral concerns. Babies, infants and actively psychotic individuals will typically not be able to respond to moral reasons *and* are typically unable to understand and meaningfully respond to forms of moral address.

One option is then to conceptualise the communication connection as a relation between insensitivity to reactive attitudes (liberally understood) and *global* incapacity to respond to moral reasons. Babies and actively psychotic individuals are (temporarily) globally incapacitated to respond to moral reasons. This global incapacity seems to robustly indicate insensitivity to reactive attitudes, and the other way around insensitivity to reactive attitudes, on a liberal reading, seems to robustly indicate global incapacity. If a person cannot in *any* way understand any form of moral address, they may be globally incapable of responding to moral reasons.

On this interpretation the case study of autism would not falsify the walk talk claim. Individuals with high functioning autism, but also individuals with obsessive-compulsive disorder, or people with kleptomania, are typically conversationally competent even though their reason-responsiveness is somewhat impaired. But all these individuals are only *locally* incapable of responding to moral reasons.

However, on this interpretation it is less clear how the communication connection is theoretically interesting. What, on this interpretation, is it that insensitivity to reactive attitudes tells us about the capacities that are required for being responsible? One may be sensitive to reactive attitudes, but not be responsible for harm due to a local incapacity. Hence a person's sensitivity to reactive attitudes does not tell us that the person has the capacities that are required for responsibility nor does it tell us what sorts of capacities are required for responsibility. And, a sceptic could still object that maybe every person can *always* resort to a local incapacity based defence; that, in our world as we know it, nobody really has the capacities that would be required for responsibility.

A further worry is that some globally incapacitated persons may in fact be relevantly responsive to reactive attitudes. One could imagine a highly emotional person who felt the sting of others anger and disapproval quite chronically but perhaps due to some disorder of impulse control or rationality is unable to reason correctly and act on those reasons. And, maybe a person who is psychotic may still have some sensitivity to reactive attitudes when this sensitivity is liberally understood. Further inquiry is needed before we can conclude that globally incapacitated persons are not at all conversationally competent. One should also ask if children are globally incapacitated, but responsive to reactive attitudes.

Maybe the ability to respond to reactive attitudes *precedes* general ability to respond to moral reasons. One possible line of further inquiry is then to investigate whether there is a developmental connection, rather than a tracking or constitutive connection, between a person's sensitivity to moral address and a person's responsiveness to moral reasons. I suggest that, on

liberal readings, a person's sensitivity to reactive attitudes indicates that there is powerful potential for this person to *develop* responsible agency (Brandenburg, 2019). Reconsidering the connection in this manner would nicely explain how individuals who are locally unable of responding to moral reasons, can communicate and –with help- come to repair or manage this inability in response to moral address. This would render the relation between sensitivity to reactive attitudes and responsiveness to moral reasons theoretically interesting, but not in the sense that it is currently proposed to be.

5.6 Conclusion

According to many Strawsonians, insensitivity to reactive attitudes is in some way connected to impaired responsiveness to moral reasons. I have argued that an analysis of high functioning autism falsifies a demanding interpretation of this 'communication connection' and urges Strawsonians to define both sensitivity to reactive attitudes and responsiveness to moral reasons in a liberal manner.

If sensitivity to reactive attitudes were demandingly understood, this communication connection would unduly imply that individuals with high functioning autism are incapable of responding to moral reasons.

This presents Strawsonians with the challenge of defining responsiveness to reactive attitudes in a sufficiently liberal way, so as to avoid unduly excluding individuals who communicate in atypical ways.

If responsiveness to moral reasons were demandingly understood the communication connection would unduly imply that a local incapacity to respond to moral reasons indicates insensitivity to reactive attitudes. Highly functioning autistic individuals can be locally incapable of responding to moral reasons. But this does not indicate they are not sensitive to reactive attitudes, when this sensitivity is liberally defined.

Hence, Strawsonians need *both* a liberal notion of sensitivity to reactive attitudes, and a liberal notion of responsiveness to moral reasons, in order to avoid unduly excluding individuals with autism from the moral community. However, on this reading sensitivity to reactive attitudes does not robustly track or constitute the capacities that render someone responsible. Both capacities may be important elements of moral agency, but they are distinct. I suggested the connection be reconsidered as a developmental one.

6 The Nurturing Dialogue

On mental disability and recognition of capacity in criminal law

Abstract

Article 12 of the United Nations Convention on the Rights of Persons with Disabilities calls for recognition of the legal capacity of all people with a mental disability in the context of civil law. When this call for recognition is extended to criminal law it seems to imply we should reject incapacity-based defences for persons with mental disabilities. But if it is difficult or impossible for a person to abide by the law because of their mental disability, it seems unfair not to take this into account. This raises the question of whether states can respond to a call for wider recognition of capacity in the context of criminal law without incurring undeserved blame and punishment of persons with a mental disability. I will argue that a modified version of the communicative theory of criminal law can provide an answer.

6.1 Introduction

Persons with mental disabilities have been, and still are, subject to a range of different sorts of abuse and discrimination. Recently there has been special concern for the lack of recognition of the capacities of persons with mental disabilities. It is harmful not to recognise a person's capacities. Among other things, such a lack of recognition provides a motive for disallowing individuals to participate in ways that they could. If, for example, we consider an elderly lady with early stage dementia to be incapable of managing her own financial affairs, this can provide grounds for appointing a guardian who will manage her affairs for her. But if this person *can* in fact participate in the management of her financial affairs, we would disadvantage her and demean her by obstructing her from such participation. Article 12 of the United Nations Convention on the Rights of Persons with Disabilities (CRPD) aims to address these types of injustices by calling for recognition of the legal capacity of all persons with disabilities, mental disabilities included (2014).

According to article 12, persons with disabilities should 'enjoy legal capacity on an equal basis with others in all aspects of life' (2014). The article also states that 'States Parties shall take appropriate measures to provide access by persons with disabilities to the support they may require in exercising their legal capacity' (2014). The CRPD hereby aims to increase recognition of the capacities of persons with mental disabilities. To illustrate, in the case of the elderly lady with early stage dementia, the state should assume that she has capacity and should take steps to

allow her to exercise this capacity. Among other things, this would mean that instead of appointing a guardian, she should be offered the assistance that she needs to make financial decisions. By recognising the capacities of persons with a mental disability, the article allows these individuals to participate meaningfully in society (2014, para. 13).

In this chapter I ask how and to what extent this call for recognition of capacity can be translated to criminal law. Criminal law, by default, assumes that an offender has certain mental capacities that would be required for being criminally responsible. Only if an offender possesses these capacities can she be considered blameworthy for the offence. An unqualified translation of the call for recognition of capacity seems to imply that, within the criminal justice system, persons with a mental disability should – just as in civil law – be considered to ‘enjoy equal capacity on an equal basis with others in all aspects of life’ (2014). This suggests that they cannot be considered unfit to stand trial, appeal to an insanity defence, or resort to other incapacity-based defences on the basis of their mental disability (for discussion see Craigie, 2015).

The problem is that such an unqualified translation would be unfair to *some* persons with a mental disability. It would be unfair to those persons with a mental disability who lack sufficient capacity to be considered blameworthy for a specific offence. This chapter asks how a criminal justice system could extend recognition of capacity to those particular individuals without unfairly blaming and punishing them.

In the next section I discuss in more detail how the criminal justice system may be capable of expressing recognition of a person’s capacities. Communicative accounts of criminal law provide the most extensive analysis of how criminal law can express such recognition. According to these accounts, sentencing should be conceived of as a moral dialogue. One only deserves to be involved in such dialogue if one is blameworthy for the offence, and meaningful participation in the criminal justice system requires capacity to participate in such dialogue. In order to be involved in such dialogue, the offender should have the capacity to abide by the law and should have communicative capacities.

The rest of the chapter is concerned with the question of how we can extend recognition to individuals who are not blameworthy due to incapacity. No communicative account has explicitly engaged with this question, but we may find the resources for an answer in how communicative theorists propose juveniles should be treated on their accounts. All communicative theorists believe juveniles should, just like adults, be engaged in a moral dialogue during the criminal justice process *even though* they are less competent than adults. This suggests that we may adopt a similar strategy for individuals with a mental disability who are not blameworthy due to incapacity. However, it is not yet insufficiently clear what sort of moral

dialogue juveniles (or anyone else) is to be engaged in when they are not blameworthy due to incapacity.

The fourth section proposes an addition to communicative accounts that would allow them to clarify how we can extend recognition to individuals who are not blameworthy due to incapacity. As well as recognising agency through a dialogue that involves the communication of blame, I suggest that one may express recognition through another type of moral dialogue that I call ‘nurturing dialogue’. A nurturing dialogue recognises communicative capacities and capacities to involve oneself in forms of moral development or management that help address past harm and future prevention. The fifth section discusses how this type of dialogue can allow for greater recognition of the capacity of individuals with a mental disability unfairly blaming and punishing them.

The sixth section anticipates an objection to my account. On Lacey and Pickard’s forgiveness-based account of criminal law, even individuals who are culpable should not be blamed. One might object that if nobody should be blamed in the first place, the addition I propose in this chapter is not needed. I explain why a nurturing dialogue would also be a relevant contribution to this particular account of criminal law.

6.2 Communicative accounts and recognition of capacity

What does it mean to recognise a person’s capacity in the context of criminal law? Communicative accounts of criminal law provide an elaborate answer to this question. Proponents of the theories I have in mind are, for example, Andrew von Hirsch, Anthony Duff, and Christopher Bennett (1996; 2003; 2008). One of their main theoretical concerns is formulating how criminal law can treat offenders in a manner that recognises their agency. According to communicative theories of law, we uniquely respect offenders as competent and morally responsible members of society by conceiving of sentencing as *a moral dialogue*. On these accounts, punishment is an expression of blame that allows an offender to receive and respond to this moral message.

This communication of blame recognises a person’s capacities in two ways. Firstly, the communication of blame recognises the offender as someone who is responsible for her crime because she *can* understand and abide by the law. If punishment would be justified solely on the basis of deterrence and incapacitation, one would conceive of (potential) offenders as people who are guided by threats and contained by brute force. One thereby does not recognise people’s abilities to govern themselves in the light of *moral* standards. Communicative theorists object that one thereby fails to respect the person as a self-governing citizen of a liberal

community. The communication of blame conceives of an offender as someone who can govern herself in the light of the standards and values of the society she lives in. As Von Hirsch points out, this does not mean that other considerations such as deterrence or crime reduction should *not* play a role in the justification of punishment at all, but respect for agency should constrain the sort of punishment practices that can be appropriate in liberal democratic societies (1996 add page). Punishment should be constrained by the communication of deserved and proportional blame. I will refer to respect for this particular type of responsible agency as respect for a person's capacity to abide by the law.

Secondly, communication of blame respects a person's capacities by recognising the offender as someone who can be (morally) reasoned with. The communication of blame aims to 'engage that person as an active participant in the process who will receive and respond to the communication, and it appeals to the other's reason and understanding – the response it seeks is one that is mediated by the other's rational grasp of its content' (A. Duff, 2003, pp. 79–80). Communication aims to engage the other as a rational human being, and when we (aim to) communicate we recognise that the other person can be so engaged. The communication of moral blame is, according to Duff, a process of rational and 'transparent' persuasion, the aim of which is to make an offender 'recognize and accept the law's requirements as being justified and refrain from crime for that reason' (A. Duff, 2003). This type of persuasion is respectful of a person's ability to participate in moral dialogue as '[i]ts aim must be to persuade them to refrain from criminal wrongdoing because they realize it is wrong' (A. Duff, 2003). Communication of blame is thus an attempt to persuade, and offers an opportunity for reform. It is ultimately up to the offender to decide if she repents or remains unmoved and unpersuaded by censure. But by allowing the offender to make this decision, one shows respect for the wrongdoer's ability to participate in a moral dialogue. This is the second type of capacity that is recognised and respected through the communication of blame. I will refer to this as respect for a person's communicative capacities.

To summarise, communicating blame to the person, on these accounts, initiates a possibly reconciliatory dialogue between the offender, the community, and (if possible) the victim. One only deserves to be engaged in such dialogue if one is found to be blameworthy for committing an offence. And engaging a person in such dialogue recognises both a person's capacity to abide by the law (which is required for blameworthiness) and a person's communicative capacities (which are required for responding to blame as a form of moral address)(Bennett, 2008; A. Duff, 2003).

Communicative accounts provide a convincing conception of how criminal law can recognise the capacities of offenders and (thereby) allow them to meaningfully participate in society. How and to what extent can and should we recognise the capacities of persons with a mental disability on these accounts? The straightforward answer is to the extent that these people *possess* both capacities, which is to say to the extent that they both possess the capacity to abide by the law and have communicative capacities.

Sometimes a mental disability does not at all diminish a person's agency to abide by the law. Susan Hayes, for example, observes that people with an intellectual disability, much to the frustration of lawyers, are a very heterogeneous group and that each case therefore requires 'a meticulous description of the nature of the offence and, if possible, the client's motivations, thoughts and feelings at the time' (Hayes, 1994, p. 289). An example she provides is that masturbating in public may be directly related to a person's intellectual disability if this person does not know the difference between public and private, but in another case the person with an intellectual disability may be well aware of this distinction and aware that his behaviour is producing the desired effect (Hayes, 1994, p. 288).

Similarly, in cases where a person has autism spectrum disorder (ASD), one has to consider the offence, the impact of the disorder, and the particular individual in order to determine whether ASD in fact undermined agency to abide by the law. Persons with autism spectrum disorder may, for example, be well aware of prohibitions against stealing and able to conform their behaviour to this awareness. But these same individuals may have trouble understanding what stalking behaviour amounts to and when a person non-verbally signals that she does not consent (Freckelton & List, 2009).

How can and should we treat individuals who are not (or hardly) blameworthy due to incapacity? Can we, in the spirit of the CRPD, somehow allow these individuals to meaningfully participate in the criminal justice process, e.g. participate in moral dialogue? And can we do so without unfairly blaming them?

As far as I am aware, communicative theorists have not explicitly asked this question regarding persons with a mental disability. There has, however, been engagement with the question of how juveniles should be treated. This issue is relevant because young offenders may not always be blameworthy due to incapacity or diminished capacity to abide by the law. How do communicative theorists believe criminal law can and should engage with these juvenile offenders?

6.3 Communicating with juveniles

In an edited volume called *Punishing Juveniles: Principles and Critique*, three proponents of a communicative approach exchange their views on how they believe juveniles should be treated within the criminal justice system (Weijers & Duff, 2002). All of them agree that juveniles should be engaged in a moral dialogue.

Lode Walgrave argues that the communicative aspect of our response to juveniles is crucial but that this communication should be more complete and rich than is possible through punishment. Adequate conditions for communication are difficult or impossible to achieve ‘in confrontation with a judge who will in the end decide upon the kind and degree of hard treatment’ (Walgrave, 2002, p. 98). According to him, a restorative setting is best suited to a communicative response to crime. The restorative justice paradigm conceives of criminal justice as an arena within which victims and the community should find reasonable reparation or compensation for harm and suffering caused by the offence. The mediation or dialogue between victims and offenders should be characterised by 1) a definition of crime as an injury, 2) an intervention oriented to restoration, 3) acceptance of the offender’s accountability and his active and direct involvement in the restorative action, and 4) the judicial framework to supervise the operations of these modes of intervention (Walgrave, 2002, p. 103). He adds that ‘participation must in no way be imposed, and agreements must be accepted and reasonable in relation to the seriousness of the harm and to the accountability and the capacities of the parties’ (Walgrave, 2002, p. 105). If the parties are unwilling to engage in dialogue or cannot come to agree, the law should resort to juridically imposed restoration.

Walgrave’s approach to criminal justice ‘focuses on harm and suffering and not on the offender’s culpability’ (Walgrave, 2002, p. 111). Walgrave therefore suggests that there need not be a radical distinction between how juveniles and adults are treated by criminal law on his account. But this conclusion raises some questions. Walgrave focuses on restoration of harm and suffering. But he also maintains that acceptance of accountability is central to the restorative dialogue. Does this mean that even non-culpable juveniles should accept accountability? If so, how would we justify this?

In response to Walgrave, Duff points out that an offender’s blameworthiness cannot be isolated from the type of restorative outcome that is required. Because, according to Duff, if an offender is blameworthy,

what needs ‘repairing’ or ‘making up for’ is not just such material harm as was caused (which might anyway be irreparable), but the *wrong* that was done to the victim; and what *that* requires is, minimally, an apology which expresses both the offender’s repentant

recognition of the wrong done and his commitment to avoid its repetition (A. Duff, 2002, p. 122).

This being said, Duff believes juveniles should be engaged in such moral dialogue. According to Duff, 'If juvenile offenders are capable of taking responsibility for their actions we should treat them as responsible agents: as agents who, inter alia, can and should be called to account for the wrongs that they do.' We should do so because we thereby 'respect them as agents who can take responsibility and out of concern for their development into fully responsible agents' (A. Duff, 2002, p. 131). But recall that according to Duff an expression of blame is only *deserved* if it is directed at a person who had capacity to abide by the law. Only if a person 'failed to recognize, accept or to be adequately motivated by reasons which were within his grasp' can the person be blameworthy (Duff and Von Hirsh 2005). Is Duff suggesting here that even juveniles who are not blameworthy but can take responsibility should be blamed in this sense?

Although Duff normally includes the resentful attitudes of victims and the moral condemnation of the community in his definition of punishment, he does not mention these particular aspects of punishment when discussing juveniles. He also states that, according to him, there is more room in the case of juveniles than in the case of adults to talk of punishment as education (A. Duff, 2002, p. 132). But on Duff's account, punishment in criminal law – by definition – involves an expression of blame (R. A. Duff, 2011). So the reader is left in the dark as to whether blame plays a role in the treatment of (all) juveniles. If it does, one wonders how this can be justified given the fact that not all juveniles will have the capacities that render them deserving of blame. If it does not, one wonders what the dialogue in which they are treated 'as' responsible agents looks like and how this dialogue would respect the capacities of juveniles.

On Weijers account, blame does play a role in the dialogue he believes juveniles should be engaged in, even if they are not (or not yet fully) responsible. He stresses how such moral dialogue has an important pedagogical role. Within the juvenile criminal justice system, this dialogue would require that the child is able to realise that the wrong he has done has social importance. If so, the young offender can be held responsible as a citizen before the law. By doing so he can and has to learn (Weijers, 2002, p. 152) what citizenship means and what being held responsible to the community over and above a parent means (Weijers, 2002, pp. 136–141).

According to Weijers, the moral dialogue through which the child is held responsible should 'attempt to get the young person to realise the moral significance of what he has done, the pain inflicted on the victim, and the harm his behaviour has done to the community and to himself as a moral being' (Weijers, 2002, p. 137). He also states that 'juvenile justice should aim at getting the juvenile offender to feel remorse for what has been done and to accept the need to

make amends to be able to restore his position in the moral community (Weijers, 2002, p. 152). On this account, juveniles should be blamed so that they can learn to take responsibility.

But on this account we may wonder whether juveniles should also be blamed if there are other – less punitive and harsh – ways in which to engage the person that would be equally (if not more) educative. Furthermore, on this account, juveniles who are not blameworthy should not be blamed if this has no clear educative effect. There is reason to doubt that blaming children who behave objectionably (generally) has a pedagogical function (Greene, 2014; Osher, Bear, Sprague, & Doyle, 2010).

It is clear that all three communicative theorists suggest that we should engage juveniles in a dialogue so as to allow them to meaningfully participate in the criminal justice system. This suggests that communicative accounts may propose something similar so as to extend recognition and meaningful participation to a greater number of persons with a mental disability.

The problem is that it is not yet clear what this dialogue does exactly and what it should look like when a person had insufficient capacity to abide by the law. In the next section I propose an addition to communicative accounts that addresses exactly this problem. My addition will, in turn, allow us to further extend recognition of persons with a mental disability within the criminal justice system.

6.4 Nurturing dialogue

How should the criminal justice system engage with an individual who is not blameworthy due to incapacity? Imagine a 10-year-old boy called Johnny with a below average IQ and an abusive parent. Johnny responds violently when one of his classmates insults him. He may have difficulty understanding the moral import of his conduct under the circumstances and may find it too difficult to control his own aggression given the stresses he is under. Johnny needs help in order to be able to understand the moral import of violence and support in order to be able to control or manage his outbursts under stress. He would not *deserve* to be engaged in a moral dialogue that expresses blame nor would he deserve to be punished in a way that expresses this blame. But the fact that an individual like Johnny does not deserve to be blamed does not necessarily mean the person cannot (and should not) be engaged in *any* type of moral dialogue. This individual may still be someone with whom we can morally reason. Recall that on communicative accounts, an engagement in moral dialogue recognises two types of capacities: capacity to abide by the law and communicative capacities. If we can find a way within the criminal justice system to recognise the individual's communicative capacities without also blaming the person, we can extend recognition of capacity without unfairly blaming them.

I propose that persons who have broken the law but are not blameworthy due to incapacity may be engaged in what I call a *nurturing moral dialogue*. So far, we have only discussed a moral dialogue that involves communication of blame and calls for a remorseful response from the offender in order to repair relationships of mutual regard. In contrast to such blaming and reconciliatory dialogues, a nurturing dialogue does not presuppose that the person is already independently able to act differently under similar circumstances. In this dialogue we suspend judgement regarding responsibility for harm and the person's 'guilty mind', because we suspect that the person is not blameworthy due to incapacity.

In a more interpersonal context, when someone acts objectionably but one doubts that this person can be said to be blameworthy, this creates the need to affirm that this person does have regard for others and the requirement for a conversation that is aimed at developing a person's capacities and/or managing impaired capacity so as to address the objectionable conduct and help prevent similar conduct from occurring in the future. The ensuing moral dialogue is one in which 1) the offender comes to recognise that (and how) someone was harmed, and 2) the parties establish forward-looking responsibilities that help address this harm by preventing similar harm from occurring.

Within criminal law, this type of dialogue would be called for when an offender is not blameworthy due to incapacity. In this nurturing dialogue, we treat the person who acted objectionably as someone who can be morally reasoned with and is, in that sense, recognised as a morally capable agent. But we do not blame or attribute blameworthiness to the person. In addition to recognising communicative capacities, this dialogue recognises that the lawbreaker can take certain future-directed responsibilities because of the harm that occurred. This is another way in which this dialogue expresses recognition of capacity and regards the person as a member of a shared community.

The appeal to the responsibilities of the offender in this dialogue relates to his or her abilities to be involved in a process of development or self-management that would prevent similar harm from occurring. If the person did not have sufficient capacity for understanding and self-control at the time of the offence, we can ask her to engage in a learning process that will help to develop this understanding and self-control, or we may ask that she tries to manage an inability to understand moral import in certain contexts. To the extent that a person can be involved in such moral development or management, she has a responsibility to do so.

It is important that this form of holding a person responsible is purely future directed. We do not hold the person responsible *for* the harm she did, but we hold her responsible for involving herself in processes of moral development and management *because* of the harm that

occurred. This form of holding someone responsible recognises the person as a responsible member of the community without also blaming the person for past conduct. Elsewhere I have referred to this form of holding responsible as the nurturing stance (Brandenburg, 2017).

I also believe it is important that in a nurturing dialogue *all* the parties should be encouraged to think about what their future-directed responsibilities are, that is, to think about how they can contribute to prevention as a way of addressing past harm. These gestures help to express affirmation of mutual regard and further stress that the offender is not held individually blameworthy *for* the harm that was done. The state, on behalf of the community, may consider how it can compensate for losses that are nobody's moral fault and may help to structure policies and institutions in ways that help prevent this type of crime from occurring. The victim may alleviate the possible guilt or anguish the lawbreaker feels by stressing that they do not hold anything against them. In this scenario, agent regret and feeling sorry *because of* the harm and hurt that has been done would of course be appropriate, but it should not be required that the offender feel guilt for what she has done. The victim may, under some circumstances, also feel a need to be minimally involved in the process that would help the offender to improve or manage in the future. Of course, none of these things can be *required* or demanded from the victim. All I mean to say here is that *if* the victim recognises that the offender had insufficient capacity, they may themselves benefit from such involvement as a way of addressing past harm and as a step towards closure.

My proposed addition can provide an answer to some of the questions that were raised about the treatment of juveniles on communicative accounts. Imagine the following case: two 10-year-old boys are playing with a ball on a high bridge. As they are throwing the ball, they widen the distance between them to increase the challenge and fun of the game. At some point, one of the boys fails to catch the ball and it bounces over the side of the bridge on to the highway. It hits a passing car and causes a serious accident. The boys, realising what they have done, run away from the scene.

If these children were to come before the court, how could and should criminal law engage them in moral dialogue? Unless there is clear evidence to the contrary, I do not think the court should assume the boys were sufficiently competent for their actions to be considered *culpably* reckless and negligent. They did not knowingly risk or harm the lives of others. Furthermore, if they had been older it may have been considered negligent of them not to call an ambulance upon witnessing the accident, but one may understand why these juveniles did not have the presence of mind to do so in this scenario.

A blaming dialogue, as proposed by Duff and Weijers, would be needlessly harsh partly

because these children do not deserve blame and partly because a nurturing dialogue can be equally – if not more – pedagogical. We can instead meet the communicative aims these philosophers are concerned with by engaging the boys in a nurturing dialogue. This dialogue displays recognition of their communicative capacities, their abilities to be involved in repair, and affirms and secures their regard for victims and the community without unfairly blaming them. By adding this dialogue to their accounts, Duff and Weijers can explain how we can recognise capacity and allow for meaningful participation when an offender is not deserving of blame due to incapacity.

What would Walgrave say about this case? It should be noted that there are victims who may be in need of some sort of restoration. Therefore, the boys should probably be engaged in a restorative dialogue on his account. If so, what does it amount to for them to ‘accept accountability’ in a way that does not hold them responsible in ways that they do not deserve? I suggest here that one can hold them responsible through engaging them in a dialogue about what should be done *because of* the harm that has occurred without also holding them responsible (and blaming them) *for* the harm that has occurred. Their acceptance of accountability is then an acceptance of some role that they can and should play in moral development or management that (helps to) prevent similar harm from occurring. Taking this responsibility helps to address the past harm by affirming mutual regard.

Of course there are a number of further questions that need to be addressed in order to explain how a nurturing dialogue could be incorporated in a criminal justice system. Among them is the question of whether such dialogue should and can take place within the court room or whether it should be a type of diversion akin to current restorative justice processes, mediation, or other alternatives. Another question is whether and how this dialogue should take place if either the victim or the offender is unwilling to participate in such dialogue. In this chapter I want focus on how incorporating nurturing dialogue, in one way or another, would allow us to extend recognition and meaningful participation to a particular group of persons with a mental disability.¹⁶

6.5 Extending recognition of capacity to persons with a mental disability

If communicative approaches were to incorporate nurturing dialogue, they could extend recognition to a particular group of persons with a mental disability who are not blameworthy due to incapacity. To be precise, they could recognise all those individuals with a mental

¹⁶ Please note that something akin to a nurturing dialogue may also be relevant in cases of moral injury sustained during war missions, moral luck cases, or automatism.

disability who can participate in moral dialogue, and can – with help – be involved in the moral development or management that is required for addressing the harm.

In the past it has been assumed that persons with a mental disability are crucially different from juveniles because they cannot develop their agency whereas juveniles can. Jeffery G. Murphy, for example, writes that in contrast to children, the class of what he calls, ‘retarded’ persons ‘is a class of persons who will never be in a position where it could reasonably be claimed that their destinies ought to be determined by their own choices and decisions’ (Murphy, 1979, p. 34). This is an example of the lack of recognition of the capacities of persons with a mental disability that has inspired formulation of the CRPD. Persons with a mild intellectual disability often have a basic understanding of right and wrong that they are able to – with help – develop and conform their behaviour to (Hayes, 1994).

The Australian Neurolaw data base reports on the case of *DPP v Bir* [2015], in which Bir was charged with causing reckless injury because he joined his male companion in the assault on the victim. The judge accepted an expert report by Dr O’Meara. The report

indicated Bir had an extremely low level of general intellectual functioning. He displayed poor performance in the areas of immediate listening attention span, working memory span, and sequencing span. His executive functioning was also severely impaired. Dr O’Meara noted that Bir demonstrated a relatively stronger ability to learn and retain information through repetition. Based on these results, she concluded Bir suffered a mild to moderate intellectual disability, and that his relative strength in repetitive learning was characteristic of male intellectual disability [42]. She also stated that she did not consider Bir’s traumatic brain injuries to contribute greatly, if at all, to the results of his cognitive testing.

Dr O’Meara opined that Bir’s intellectual disability and his excessive alcohol consumption contributed jointly to his offending [46]-[47]. She considered that Bir’s poor reasoning and problem-solving abilities, as well as his slowed information processing, would have affected his capacity to comprehend situations and make appropriate decisions. She regarded Bir as highly susceptible to external influence, and observed that whilst Bir was able to comprehend the wrongfulness of his actions in clear-cut situations, he likely struggled with processing wrongfulness in fast-paced scenarios. Dr O’Meara opined that at the time of offending, these pre-existing cognitive impairments would have been exacerbated by Bir’s state of intoxication (2015).

On the basis of this report, one may strongly suspect that the person’s agency to abide by the law did not meet the threshold that would be required for deserved blame. Nevertheless, the fact that Bir has a basic grasp of the wrongfulness of his actions in clear-cut situations and a capacity

to learn suggest that he may be an apt candidate for the nurturing dialogue discussed above.

Because Bir has this understanding, he can affirm the wrongfulness of his past conduct (feel sorry about what happened) and can affirm mutual regard by taking certain forward-looking responsibilities that help address this harm. He can also be asked to take certain responsibilities within the process that would work towards restoring some of this harm and preventing future harm. That is to say that we can and should ask of Bir that he involve himself in a learning process or process of self-management that would stop similar harm from reoccurring.

If the person in question can participate in such dialogue, it would be problematic to obstruct him from participating in it on the basis of his below-threshold agency to abide by the law. It would be problematic because this obstruction expresses misrecognition of his (other) capacities and thereby demeans him, and furthermore because he is denied the possible benefits of this dialogue: an opportunity to find closure, affirm regard, and be involved in a process of self-correction and development that contributes to restoration and helps prevent future harm.

This proposal for extending recognition of capacity is also especially relevant for criminal justice regarding persons with ASD. Such persons will often be able to participate in moral dialogue and will be able to develop their agency even when their agency to abide by the law fell below the threshold at the time of the offence. Freckelton discusses how ASD may impact upon the capacity of a defendant to form the necessary intent for criminal offences (Freckelton & List, 2009). A case that illustrates this impact is the case of *Parish v DPP* (2007). Mr Parish behaved objectionably towards a woman on the train:

Very soon after the train started its journey, Mr Parish pushed his calf against hers. She tried to move her leg away from his. He also changed positions and sat directly in front of her, with his hands over his knees. He then placed his hands on top of her knees. At this stage she was looking out the window, trying to ignore him. He then rubbed his hands on top of her knees. She did not speak to him or attempt to change seats. When she got to Box Hill Station, she waited back and allowed Mr Parish to alight first. She then stood beside the train to make sure that Mr Parish was away from her. She then proceeded to take the escalator but felt a hand on top of her hand. She turned around and noticed that it was Mr Parish. He was standing on the step below her and, as the escalator was going up, he rubbed her lower back and her upper buttocks. She gave evidence that she was scared and was unable to move through the people surrounding her on the escalators. Again she said nothing to him. After alighting from the escalator, she went to look for her sister and broke down in tears when telling her what had happened. On the recommendation of family members, she made a complaint to the police. Mr Parish was identified from CCTV photographs (Freckelton & List, 2009; *Parish v DPP*,

2007).

According to an expert report, Mr Parish, because he had Asperger's syndrome, 'would have been unlikely to have been aware that the complainant was not consenting to his actions' (Freckelton & List, 2009, p. 262). After the appeal, Mr Parish was found not guilty on all charges of assault. Justice Robson did, however, stress that he 'would expect that in the case of a person who was not suffering from Asperger's Syndrome or having a similar disability, that the prosecution would be able to easily establish the necessary awareness on the part of any person who did what Mr Parish did' (Freckelton & List, 2009; Parish v DPP, 2007).

This kind of case lends itself to the creation of a nurturing dialogue between the victim, the community, and the offender that, among other things, expresses recognition of the person with ASD as a person who can have regard for others and can take responsibilities that are related to expectations of mutual regard. This dialogue would not involve blame but it would aim to come to an understanding about the wrongdoing and to appeal to the offender to involve himself in a process that would prevent future objectionable conduct. Restoration may consist of the wrongdoer taking steps to learn how to refrain from inappropriate sexual behaviour in the future. We can also imagine that the wrongdoer may offer further contributions to future reform out of his own free will, such as contributing to raising awareness among other persons with ASD. The victim may plausibly also be concerned about raising such awareness, and may even want to assist. The exact agreed-upon future-directed responsibilities will depend on the dialogue that takes place. But this case illustrates how it is conceivable for a victim to feel the need to be involved in a process that affirms mutual regard and works towards future prevention. Of course, state parties should in a case like this be concerned about the lack of structural education or help provided to individuals with high-functioning autism when it comes to navigating the social world, which includes registering and responding to forms of non-verbal consent. A dialogue between the three parties would bring out these shared responsibilities that flow from an affirmation of mutual regard.

I have illustrated above how nurturing dialogue may allow for further inclusion of a group of persons with a mental disability in the criminal justice process without thereby unfairly blaming them. The task ahead of us is to look into whether and how persons with a mental disability whose agency to abide by the law falls below the threshold *can* nevertheless participate in a nurturing dialogue and can – with help – be involved in the development or management of their own capacity to abide by the law. This involves asking what are adequate conditions for dialogue; what form or style of conversation the person is receptive to; and how much responsibility one minimally should, and maximally may, ask the person to take. In so far as there

is a way to engage the wrongdoer in a nurturing dialogue, we can and should recognise their capacities by providing this opportunity for dialogue to the person and to the victim.

At this point, we are also able to translate the third requirement of article 12 to criminal law. Recall that article 12 of the CRPD, among other things, states: ‘States Parties shall take appropriate measures to provide access by persons with disabilities to the support they may require in exercising their legal capacity’ (2014). The introduction of nurturing dialogue allows us to translate this third requirement to the context of criminal law. Persons with mental disabilities should be offered assistance so as to allow them, first, to exercise their communicative capacities within the context of criminal law and, second, to involve themselves in developing or managing their capacity to abide by the law. One may support communicative capacities by offering assistance with regard to the sort of environment in which a dialogue takes place, the people that are present, and the ways in which a person may express herself. Facilitating a person’s ability to involve themselves in developing or managing their capacity to abide by the law will rely on measures that are similar to allowing people to be agentially involved in their own recovery. In the legal context this would more specifically amount to exploring the ways in which persons with mental disabilities can have agency to repair or manage their own capacities to abide by the law.

To conclude, in the context of criminal law we can further extend recognition of persons with a mental disability without unfairly blaming them by recognising and supporting their capacity to participate in a nurturing dialogue. A nurturing dialogue expresses recognition of a person’s capacity to communicate and of a person’s capacity to participate in moral development or management. This dialogue allows for an extension of recognition of capacity and enablement of meaningful participation for a larger number of persons with a mental disability than our current criminal justice practice does without thereby incurring unfair blame and punishment.

6.6 Anticipating an objection

Lacey and Pickard propose that sentencing should not be conceived of as communication of blame but, instead, as communication of forgiveness (Lacey & Pickard, 2015). On blame-based accounts, forgiveness and reconciliation are a *possible* outcome of a moral dialogue that communicates blame. But Lacey and Pickard propose that we forswear affective blame entirely in our communication with the offender and conceive of punishment *as* forgiveness (Lacey & Pickard, 2015, p. 677). On their account, punishment is not in any sense retaliatory. Instead punishment should be conceived of as reparative costs and a risk-reductive strategy that is a *consequence* of responsible wrongdoing (Lacey & Pickard, 2015, p. 678). What we communicate by

sentencing is that we forgive the person for what she has done. Importantly, communicating forgiveness presupposes responsibility for wrongdoing and as such still recognises an offender's capacity to abide by the law. We subject the person to reparative and risk-reductive strategies because she *chose* to offend. Punishment as forgiveness affirms the responsibility of the person but rejects the idea that this responsibility provides sufficient ground for affective blame and retaliation (Lacey & Pickard, 2015).

One may wonder whether punishment as forgiveness leaves sufficient agency for the offender to respond in ways that would merit forgiveness. Blame-based accounts provide an offender with an opportunity for reconciliation if they *decide* to show remorse and make amends. But on this account the offender's response is not a condition or reason for forgiving the person (compare also Milam, n.d.). One may also wonder if the state should be allowed to forgive on behalf of victims and communities. Nevertheless, this proposal provides a unique way to affirm a person's responsibility without also blaming the person. There is merit to this proposal because blaming is a harsh form of treatment that may be needlessly harsh or fail to rehabilitate. Let us, for the sake of argument, accept that criminal law should ideally provide punishment as a path to forgiveness. Would this render the need for nurturing dialogue redundant?

One might think that nurturing dialogue becomes redundant, because on a forgiveness-based account no one will be unfairly blamed. The nurturing dialogue aims to extend recognition of capacity without incurring unfair blame. Hence, if no one will be unfairly blamed, this addition may no longer be required. But, as I argue below, punishment as forgiveness at most reduces the scope of nurturing dialogue but does not render these types of dialogue entirely redundant. It does not render them redundant, because forgiving a person can, under some conditions, also be unfair and forgiving a person should, under other conditions, be supplemented with a nurturing dialogue.

Conceiving of punishment as forgiveness may make for a dialogue that is appropriate for a number of offenders whose capacities are impaired. The reason for this is that the threshold capacity required for deserved blame and punishment in the legal sense may arguably be higher than the threshold capacity that would be required for punishment as forgiveness. For example, one may consider Bir, whose offence was discussed above, to be somewhat competent and responsible but not sufficiently competent to be worthy of blame and punishment. However, once we conceive of punishment as a path to forgiveness, our intuition may be that Bir is sufficiently competent to merit being subjected to this type of communicative punishment. So punishment as forgiveness may be considered appropriate in a case where punishment as blame would not be. Consider also how Pickard and Lacey are often specifically concerned with

individuals with borderline personality disorder or an addiction who end up in court (Lacey & Pickard, 2013b). The majority of these individuals will have *some* level of control, but the level is likely to be significantly lower than average. One could argue that this level of control suffices for punishment as forgiveness but not for punishment as blame.

However, although a forgiveness-based dialogue may be appropriate for some individuals who would not merit being blamed due to compromised capacity, it does not thereby dissolve the need for nurturing dialogue. First, to the extent that it is difficult for the person to independently act differently in similar circumstances due to incapacity, forgiveness should still be supplemented with a nurturing dialogue. The person should in these cases not ‘just be forgiven’ but should also be engaged in a dialogue that addresses how capacity may be developed or managed and should be asked how they – with help – can be agentially involved in this process.

Second, there will also be cases where a forgiveness-based dialogue is entirely unfair and the person should only be engaged in a nurturing dialogue. In cases where a person is not on any level responsible *for* harm, punishment as forgiveness would still amount to unfair treatment of the person. First, it would be unfair because punishment as forgiveness attributes responsibility for harm to the person, which would in this case be erroneous. Imagine, for example, that Alisha arrived late to an appointment because the bus she was on broke down and she had to wait for a replacement. If Patrick, who was waiting for her, were to say to her, ‘I forgive you for making me wait’, this would unfairly attribute responsibility to Alisha for arriving late. Miranda Fricker has pointed out that, in addition to this, by forgiving instead of blaming one would render it difficult for Alisha to defend herself in response to this allegation because Patrick has already decided not to hold it against her (Fricker, 2018). In the legal context, Lacey and Pickard’s account of punishment as forgiveness would be equally problematic when it attributes responsibility for harm to a person who is not responsible for harm. And punishment that is a path to forgiveness would also involve unpleasant consequences of responsible wrongdoing that a person does not deserve when she is not responsible.

To illustrate, think of the case of Mr Parish discussed above. If he has no capacity to independently understand that the other person may not consent under these types of circumstances, it could be considered unfair to subject him to punishment as forgiveness. This form of punishment would convey to Mr Parish that he had *some* awareness of the risk or possibility that the other person was not consenting, and is therefore charged with assault, but that he will be forgiven after he has paid his due. Mr Parish may feel unfairly treated by this because he meant no offence and cannot grasp how he failed to regard the other person’s

interests and concerns. Furthermore, punishment for assault is still an unpleasant consequence that he does not deserve. He does not deserve to carry reparative burdens, and it seems wrong that he would be punished so as to reduce the risk that he will choose to reoffend, because he did not choose to offend in the first place. A nurturing dialogue would in this case be more likely to secure relationships of mutual regard between the victim, the offender, and the community.

Conclusion

In this chapter I asked how one could translate the call for recognition of capacity of persons with a mental disability from civil law to criminal law without unfairly blaming and punishing them. I argued that individuals who are not blameworthy due to incapacity might still have their capacities and responsibilities recognised by being engaged in a nurturing dialogue. We can recognise and respect the agency of these individuals by recognising that they can be reasoned with and can take responsibility by involving themselves in processes that should help prevent future harms. Doing so helps to address the past harm and affirms relationships of mutual regard.

One might object that on an account of punishment as forgiveness, there would be no need for nurturing dialogue because nobody would be unfairly blamed. But forgiveness can also be an unfair response to persons who are not responsible for harm due to incapacity, because, even if it is fair, it may still need to be supplemented with a nurturing dialogue for those individuals whose capacities are compromised. Therefore, a nurturing dialogue would still be needed in order to fairly recognise and include those particular individuals as persons who can meaningfully participate in the criminal justice system and, by extension, in a liberal community.

When it comes to extending recognition of capacity to persons with a mental disability within criminal law, without unfairly blaming (or forgiving) those persons, the way forward is to ask how we can recognise and support their abilities to participate in a nurturing dialogue.

7 Detached Blame and Nurturing Reproach

Therapeutic responses to objectionable conduct by service users in mental health care

Co-authored with dr. D.W. Strijbos

Abstract

In this chapter we discuss two different ways of holding service users responsible in mental health-care contexts. Hannah Pickard has provided an account of how service users should be held responsible for objectionable conduct within contexts of mental health care. Her account called ‘the clinical stance’ is based on what are considered best practices in the treatment of people with PD. From a clinical stance one holds a person responsible for harm, but refrains from emotionally blaming the person and only blames in a detached sense by considering the person responsible for objectionable conduct and resulting harm. In this chapter we compare this stance to what are considered best practices at a specialized clinic for people with ASD in the Netherlands. We refer to our alternative account of best practices as ‘nurturing reproach’ and discuss relevant similarities and differences between detached blame and nurturing reproach. We conclude with suggestions for further research and theorizing.

7.1 Introduction

Should we hold a person responsible for morally objectionable conduct if this person's ability to abide by the related moral norms is compromised? In philosophical theory there has recently been a surge in engagement with this question. The answer will of course depend on a number of factors. How and why is the norm-transgressor's ability compromised? What is the nature of the norm in question? What is the relationship between the person who is holding responsible and the norm-transgressor? And so on. In this chapter we focus on responsibility practices within contexts of mental health care. The question to be addressed is then more specifically whether, and if so how, clinicians should hold service users responsible for transgressing a moral norm when mental illness compromises the service user's agency to abide by this norm.

Hannah Pickard (2011, 2013) has developed a normative account of how service users with personality disorders (PD) should be held responsible for objectionable conduct in mental health care settings. Her account is based on what she considers best practices in a variety of effective psychological treatments of PD¹⁷. She argues clinicians should ideally take, what she calls, a *clinical stance* towards service users who transgress a norm. From a clinical stance, one holds a person responsible for transgressing a norm but does not feel and express blame related emotions towards the person. One would, from this stance, only privately and abstractly judge that a person is blameworthy for what she did, without experiencing accompanying feelings of indignation or resentment.

In this chapter we test Pickard's account, by comparing it to an account of best practices at a specialized clinic for people with Autism Spectrum Disorders (ASD) in the Netherlands. We base our account on a small qualitative study in which we compared Pickard's account to what are considered desirable responses to moral norm-transgressions in this particular clinic. We call our alternative account of best practices 'nurturing reproach'.

Both nurturing reproach and the clinical stance recommend a response that finds a mean between condoning and blaming, but they do so in a different manner and for different reasons. We discuss what lessons can be drawn from this comparison and provide suggestions for further research and theorizing about responsibility practices in psychiatric settings.

¹⁷ In her (2011) Pickard mentions varieties of cognitive behavioural therapy (DBT, STEPPS), motivational interviewing techniques, MBT, and therapeutic communities

7.2 The clinical stance

Hannah Pickard (2013, see also 2011) has recently provided a normative account of responsibility practices in contexts of psychiatric care. As a novice clinician, she observed that the ways in which patients were held responsible in a therapeutic community for people with PD were very different from what she was used to outside of this context. In this clinic ‘Service users were responsible for their actions and omissions and accountable to the Community for them, but an attitude of compassion and empathy prevailed, and they were not blamed.’ (Hanna Pickard, 2013, p. 1135)

According to Pickard such a stance is generally desirable in clinical contexts where service users suffer from a so-called ‘disorder of agency’. Pickard writes that the ‘core diagnostic symptoms or maintaining factors of disorders of agency are actions and omissions: patterns of behavior central to the nature or maintenance of the condition (2013, p. 1135).’ She mentions PD, substance use disorders and eating disorders as examples of disorders of agency. The behaviors that are constitutive of these disorders (e.g., deliberate self-harm, reckless and impulsive behavior, attempted suicide in Borderline PD, drug consumption in SUD, eating too much or too little in eating disorders) are not merely bodily movements. Rather, Pickard argues, these behaviors often constitute voluntary actions, by which she means ‘that the agent can exercise choice and at least a degree of control over the behavior’ (2011, p. 212). As evidence for this claim, she observes that on most occasions, service users who display these kinds of behavior routinely choose to behave otherwise when they have an incentive and are genuinely motivated to do so.

When a person’s agency is in this way disordered, she believes it to be important that service users are held responsible for problematic patterns of behaviour. More specific: ‘so long as they know what they are doing, PD service users are responsible for their behavior to the degree that can exercise choice and control over it’ (2011, p. 213).¹⁸ These behaviours are among the main reasons for the person seeking or needing treatment and those behaviours have to be addressed for the person to recover. According to Pickard these behaviours are best addressed by holding the person accountable for objectionable actions and omissions, thereby augmenting the service users’ existing capacity for agency – an essential part of effective treatments. But at

¹⁸ Pickard mentions two important caveats. “First, service users with PD may not always have full conscious knowledge of why they are behaving as they do, or what the full effects of their behavior on others may be [...] Second, it is important to recognize that, on the common sense conception of agency presented above, control is a graded notion, and the degree of control possessed by PD service users may sometimes be diminished compared with the norm.” (ibid.)

the same time one should not blame a person, because blaming would be highly detrimental to therapy as it may trigger feelings of rejection, anger, and self-blame, which undermine recovery.

Holding service users responsible in this sense may involve ‘asking them to explain why they made the choices they did, and encouraging them to behave differently in the future. Alternatively, it may involve the agreed imposition of negative consequences, to increase motivation, and show that the behaviour, and the harm it causes, is taken seriously.’ (2013, p. 1141) These forms of holding responsible avoid the so-called ‘rescue trap’, a response where the person is solely seen as a victim of her disease who cannot help doing what she does (Hanna Pickard, 2013, p. 1138). We would like to add here that these types of interactions are also distinct from condoning norm-transgressive conduct. If one condones problematic behaviour, one does not hold a person to account for what she did either. Contrary to rescuing, condoning does not amount to any type of intervention in the person’s behaviour. Both rescuing and condoning arguably enable rather than address disorders of agency.

The clinical stance avoids forms of rescuing or condoning by holding the person responsible for her conduct. It does so without blaming the person for what she did or does. Blaming a person, according to Pickard, is not compatible with the therapeutic relationship. Pickard has a particular account of what a blaming response exactly consists in. She discusses ‘reactive attitudes’ as the sort of reaction that amounts to blame. Philosophers have different definitions of reactive attitudes, but these attitudes at least minimally refer to those emotional responses we have towards perceived morally objectionable conduct (e.g. Deigh, 2011; McGeer, 2011; Shoemaker, 2015; Wallace, 1996). Typical examples are resentment and indignation. Pickard has her own specific definition of these attitudes. She argues that the type of hostile anger that amounts to blame always comes with a sense of entitlement: if someone is blaming ‘they feel that the other is blameworthy and so deserves their anger.’ According to her these feelings can come apart from our considered judgments (Pickard, 2013). We would like to flag here that the question of what blame-related emotions exist in is contested, and whether blame need be emotional in kind is contested too. The importance of this disagreement is brought into sharp relief when we look at the concrete examples of blaming that Pickard discusses in here online learning module. Here blaming reactions are typically described as the person being cold and dismissive combined with an insinuation that that the person withdraws from the relationship. One could explain this difference by assuming that Pickard conceives of these reactions as one important way in which people tend to give expression to blame related emotions. We will return to some alternative definitions of blame later in the chapter, and stick with Pickard’s definition for now.

An example of what a clinical stance in response to morally objectionable conduct amounts to is provided in Pickard's online module for people who work in mental health care (Pickard 2018). Two police officers are called to check on Amy, a person with personality disorder who is at that point in time threatening to people in her hallway and appears to be under influence. Amy has to be taken to hospital for a check up but once at the hospital Amy becomes violent whilst waiting for a doctor. She assaults a police officer and is consequently arrested. Later at the police station Sarra, one of the police officers, walks into Amy's cell with a cup of tea.

At this point Pickard discusses three responses to Sarra. From a rescuing response, Sarra brings Amy a cup of tea and is trying to comfort Amy and tells her she will be the one who is going to interview her. Sarra then gets angry and throws the cup of tea through the room screaming that she will file a complaint. A blaming response is illustrated by a cold reply 'go ahead and do that', and by slamming the door when leaving. Furthermore, in the blame scenario, the police officer describes her to others as 'a pain' and difficult to deal with. From a clinical stance Amy is told that she is free to file a complaint, that this is her choice, but that she will also be interviewed now for her earlier assault of a police officer. This is told to her without anger and in a 'matter-of-fact like manner'.

According to Pickard the first rescuing response sees Amy too much as a victim of her disorder who cannot help acting as she does whereas the blaming response is counterproductive to Amy's recovery. From a clinical stance one would consider Amy responsible for her earlier assault and current aggressive conduct without also blaming her for engaging in such objectionable behaviour. Therefore, some of the above manners of 'holding a person responsible' may apply. Amy has to bear the consequences of being aggressive and verbally abusive towards others, and it was her choice to drink too much. Amy is also charged with assaulting a police officer; she will be interviewed about this and possibly fined. In response to her abusive conduct the officer may stress some of those responsibilities and choices that Amy has. This last response of responsibility without blame is seen as the therapeutically effective response.

Therapeutically effective as the clinical stance may be, this stance initially seemed conceptually confusing to Pickard. For, how can one hold a person responsible for objectionable conduct, without also blaming the person for objectionable conduct? If Amy is to be held responsible for throwing a cup of hot tea through the room, why should one not also blame her for such a rash action that may end up getting someone seriously hurt? In accordance with most philosophical

theories, and arguably common sense too, being responsible for objectionable conduct *implies* being blameworthy for this objectionable conduct and the resulting harms too.

One could object here that maybe Amy is only causally responsible. Causal responsibility can come apart from blameworthiness. Though Amy caused the breaking of the cup, she was not in the right state of mind to act responsibly. However, in such a case one is left to wonder why one may appropriately hold the person responsible at all. We need a conceptual framework that explains why the person should not be blamed, but *is* responsible in a way that renders it appropriate for the person to be ‘held responsible’ in the sense described above.

In order to make sense of responsibility without blame Pickard distinguishes between two types of blame: detached blame and affective blame. Service users can be appropriately held responsible for their conduct because they are blameworthy. Detached blame amounts to the non-emotional judgment that the service user is blameworthy for what she did. As we saw, Pickard thinks that service users who are held responsible are not fully excused from blameworthiness for the norm-transgressions they engaged in because they can exercise choice and a sufficient degree of control over their behaviour. Surely for people with mental illness this degree of control is often lower than average, and their circumstances mitigate their blameworthiness. According to Pickard these circumstances need not suffice for being excused (2013, p. 1142). In a detached sense, considering the person blameworthy and responsible for what she did would therefore be a correct evaluation of this person. This mere judgement that a person is blameworthy, e.g. detached blame, is compatible with a clinical stance. Affective blame is what is detrimental to therapy and is what one, from a clinical stance, will not engage in (2013, pp. 1142–1146). These feelings and sense of entitlement often accompanies judgments of blame, but they do not necessarily do so.

As discussed above, affective blame refers to the feelings of resentment, indignation or anger combined with a sense of entitlement to these feelings. On Pickard’s account responsibility without blame specifically refers to responsibility without affective blame. This solves the puzzle of how it can be appropriate to hold a service user responsible for what she did within a context of psychiatric care whilst explaining why it is inappropriate to blame a person at the same time. The person is blameworthy to the extent that she is responsible, but engaging in affective blame is inappropriate as it undermines recovery.

We are now in a position to list a few defining elements of the clinical stance. The clinical stance recommends a response of detached blame, that:

- 1) Is appropriate given the person’s sufficient ability for choice and control and (mitigated) blameworthiness for what she did.

- 2) Is *not* accompanied by felt and expressed affective blame towards the person
- 3) Avoids a condoning or rescuing response to the person

This clinical stance, according to Pickard, fosters therapeutic recovery and as such provides a rough ethical guideline for responsibility practices in contexts of psychiatric care where it concerns service users for whom disorders of agency are (partly) constitutive of the disorder.

7.3 Nurturing reproach

In the following sections, we assess Pickard's clinical stance by comparing it to what are considered best practices in response to objectionable conduct in another mental health care setting. We interviewed clinicians at a clinical centre in the Netherlands specializing in the inpatient treatment of adults with (co-morbid) autism spectrum disorder with an average or above-average IQ. The centre provides so-called tertiary care for service users whose treatment in regular psychiatric inpatient or outpatient settings has proved ineffective. Most of the service users admitted met the criteria for one or more co-morbid diagnosis such as PTSD, depression, anxiety disorders, eating disorders, addiction, AD(H)D, or personality disorder. At the time of the study, they were 18 to 45 years old, with an average age of about 25 to 30. The male/female ratio was approximately 30/70. A significant subgroup of the service users admitted had been diagnosed with ASD at a relatively late stage of their development, often in early adulthood. Typically, these people had received treatment directed at co-morbid (sometimes incorrect) diagnoses in the past, such as depression, eating disorder, addiction or personality disorder, and had had a long, unfortunate and sometimes adverse history in mental health care characterized by a chronic misunderstanding of and by others, aloneness, growing despair, distrustfulness and alienation.

Many, if not all of the service users admitted manifested at least some degree of 'disorder of agency', in Pickard's sense explained in the previous section. Destructive or otherwise debilitating patterns of behaviour (e.g., verbal or physical aggression towards others, breaking furniture or cutlery, deliberate self harm (cutting, head banging, strangulation, self-embedding), restrictive eating and drinking, running away, suicide attempts) and extreme (social) avoidance or inactivity were among the central maintaining factors in the chronic problems experienced by the service users (and their families), severely undermining the possibility of recovery. Moreover, the challenging behaviours displayed were not 'mere movements', but on many occasions allowed for a minimal possibility for choice and control. Here, the caveats Pickard mentions (see footnote 3) are especially important. Due to problems on the level of social interaction and communication (a core diagnostic feature of ASD) and underlying socio-cognitive capacities

relating to empathy and ‘theory of mind’, many service users had serious difficulties realizing why they were behaving the way they did and what the consequences of their behaviour may be, for themselves and others.¹⁹ Also, the degree of control over their behaviour fluctuated and was significantly limited, against the background of immature coping mechanisms, rigidity and habit formation, problems with executive functioning and structural difficulties in signalling emotions and distress.

The clinicians interviewed (see below) had serious doubts or simply did not know whether these service users had the capacity of choice and control as suggested in Pickard’s writings on PD. In some cases, they were positive that the capacity for choice and control was significantly more compromised than described there. Nevertheless, the considerations Pickard discusses in relation to PD seem to at least partly apply to these service users with ASD. The core of treatment consisted in helping them to better understand the way ASD ‘works’ in their lives, how to address their basic needs and impairments in light of this using the skills they have, and how to alter maladaptive coping patterns that stand in the way of recovery. The basic assumption behind such a treatment program is, as Pickard put it, that therapy can augment existing agential capacities, capacities that, though limited, *are* present in the problematic patterns of behaviour displayed.

In an exploratory study, the first author conducted 10 semi-structured interviews with clinicians. The participants were four psychiatrists, two senior clinical nurses, one social care worker, one clinical nurse, one family therapist, one psychologist/remedial educationalist and one clinical nurse specialist, and consisted of four men and seven women. The participants were invited by email for an interview. Of the invited participants, no one refused to be interviewed. The interviews were audio-recorded and took place in a quiet interview room at the clinicians’ place of employment. They lasted between 60 and 90 minutes.

We interviewed clinicians about what they considered desirable responses to challenging behaviour.²⁰ We focused on responses to behaviour that the interviewed clinicians perceived as

¹⁹ For a recent overview of neurocognitive characteristics of adults with ASD, see Geurts and Lever (2017).

²⁰ A definition of challenging behaviour often used in the clinical literature is the following: “culturally abnormal behaviour of such an intensity, frequency or duration that the physical safety of the person or others is likely to be placed in serious jeopardy, or behaviour which is likely to seriously limit use of, or result in the person being denied access to, ordinary community facilities” (Emerson, 2001). The concept of challenging behaviour is a socially constructed, descriptive concept. What behaviour is and is not considered ‘challenging’ is subject to various sociocultural (moral) norms, service delivery patterns and changes over time and across different (cultural) contexts (Emerson, 2001; Xeniditis, Russell, & Murphy, 2001). The concept carries no specific diagnostic significance and carries no aetiological implications per se. It covers a diversity of behavioural phenomena across a heterogeneous group of people. Challenging

hurting, disrespecting or endangering other people (the clinicians themselves and their colleagues included).²¹ The interviews were semi-structured. Among other things, the clinicians were asked if they had emotional responses to this objectionable conduct, considered the person blameworthy for this conduct, and held the person responsible. They were also asked what they considered to be the desirable response and why, and were asked if they would respond differently in non-psychiatric contexts.

The derived account of best practices was presented to the interviewed clinicians in a focus group, in order to confirm and calibrate. The focus group was attended by eight of the participants: three psychiatrists, two senior nurses, one family therapist, a nurse and the psychologist/remedial educationalist. All participants were informed of the data analysis and proposed account of best practices either during the focus group meeting or by email. All participants signed a statement of informed consent and their answers will remain anonymous.

There are three common elements to what were considered the right responses to objectionable conduct in this particular mental health care context. First, when one perceives the service user to behave objectionably towards others and thereby cross the boundaries of what can be tolerated within the clinic (from now on we refer to such conduct as objectionable conduct), the clinician will typically be negatively affected by such conduct and it is considered desirable to set clear boundaries and firmly correct the service user for such conduct. In doing so the clinicians would often communicate how they were affected by this conduct and why. Negative emotions here include (mild) anger, but also shock or fear.

Second, one should correct the service user in a manner that does not threaten the therapeutic relationship and is consistent with an overarching concern for therapeutic recovery. This means that the affective attitudes of the clinician should not be communicated in an excessive and unfiltered manner, and that it should be clear this response does not amount to rejecting the person or breaking the therapeutic relationship, over and above a negative evaluation of this person's conduct. One clinician put this succinctly when she said: *'The patient should know this does not mean I will not see you anymore'; 'the possibility for contact is never foreclosed'*. Furthermore, concern for therapeutic recovery and maintenance of the therapeutic relationship, means the affective response should be attuned to the person's ability to understand and respond to it in a meaningful and constructive way. Within a clinical population of people with (co-)

behaviour may, but need not, be related to intellectual disabilities or psychiatric conditions, either as a primary or as a secondary manifestation of the condition (ibid.).

²¹ We here limit ourselves to other directed behaviour because reactive attitudes are understood to be a response to conduct that is objectionable in its treatment of other people, and more generally philosophical theorizing about blame and responsibility typically focuses on socio-moral norms and socio-moral transgressions.

morbid ASD, this required, for example, careful assessment of the service users' level of social-emotional development, their understanding of the norms transgressed, their particular and sometimes idiosyncratic ways of making sense of the (social) world and current levels of distress. The specific form of nurturing reproach that is desirable will therefore vary depending on the person, and the context.

Third, the clinicians typically did not attribute blameworthiness or sufficient choice and control to the patients. When we asked clinicians to explain and justify their responses, they did not mention blameworthiness as providing a consideration in favour of correcting the person. When explicitly asked if they considered any of the service users whom behaved objectionably to be blameworthy for what they did, most interviewees replied in the negative and some said they did not know. We also specifically asked if they considered these service users to possess the level of agency that would suffice for blameworthiness. In answer to this question, the clinicians would often report that they did not know, or could not know given the range of difficulties that may compromise the person's understanding or control in different contexts and at different moments in time. Sometimes they reported to be quite certain that the service-user did not possess the required level of understanding and control. The clinicians did consider service-users responsible in some sense but they were not considered to be responsible *for* the objectionable conduct and related harms. The clinicians attributed other responsibilities to service-users, like a responsibility to participate in shared therapeutic processes, or a responsibility to respond to nurturing reproach.

We refer to the response that is characterized by these three elements as 'nurturing reproach'. This response is an expression of a stance referred to in the literature as 'the nurturing stance'. From this stance one relates to the other person as someone who cannot yet sufficiently live up to certain interpersonal norms, but who is sensitive to moral appeal and capable of moral development (Brandenburg 2017). We will say more about this in the last section of the chapter.

Our clinician's responses were typically charged with some degree of negative affect, while also being consistent with therapeutic aims and considered 'attuned to' the service user's cognitive and social-emotional profile. These affective responses were informed by or moderated by therapeutic aims of recovery, enabling the service user to 'grow' in applying self-governing capacities conducive to recovery. Words like 'reproof', 'reproach', 'reprimand' or simply 'firm correction', or 'disapproval' could be considered apt to describe these attitudes, depending on specific context and persons involved. We will use the term 'nurturing reproach' as a general concept that may refer to any of these emotional responses. Although nurturing reproach is

characterized by negative affect, we do *not* mean for this affect to refer to specifically blaming emotions. As we will discuss later in the chapter, these clinicians reported to not attribute blame or blameworthiness to the person. Rather, nurturing reproach here indicates that the person may be negatively affected by the perceived objectionable conduct and that there can be a place for mild forms of negative affect within a clinical setting which ensue in a form of holding the service user to account in a manner consistent with therapeutic aims. Nurturing reproach does not result from being angry *with* a person but rather results from being angered by objectionable conduct²². According to the clinicians we interviewed, having some such negative affective response to objectionable conduct towards others shows involvement with the service user as opposed to detachment. This affect need not necessarily be expressed in how one corrects the person but it would be wrong to actively hide or obfuscate that one *is* affected when one communicates disapproval of the service user's conduct.

The interviewed clinicians stressed that adopting a stance from which one is not at all negatively affected by objectionable conduct towards others, or actively hides or obfuscates these felt responses, would undermine the therapeutic relationship and the service user's recovery. But why is nurturing reproach considered the desirable response when the clinician perceives a patient's conduct towards others to be objectionable? The clinicians converged on a number of reasons that according to them speak in favour of this response to perceived objectionable conduct.

1. Congruence

Part of a successful therapeutic relationship is that clinicians behave in a genuine way towards service users and that they *remain* themselves when the therapeutic relation comes under pressure in the face of the service user's challenging behaviour. Communicating one's corrective attitudes in response to perceived morally problematic conduct makes for a personal and transparent way of relating to the addressee. This genuine and clear form of relating to a person under these circumstances is closely connected to the Rogerian principle of congruence (Rogers, 1957). A congruent therapist is a therapist who is genuine, involved

²² It should be noted that we focus on responses to objectionable behaviour towards others. Within this moral domain, a transgression is especially likely to trigger negative affect. Not all of the examples of objectionable behaviour Pickard discusses fall clearly within this domain. When they don't we are not sure nurturing reproach would be a desirable response. A person self-harming or not taking medicine, for example, need not similarly negatively affect caretakers, as these are not clearly forms of hurting, disrespecting or endangering *others*. Similarly in other care-relationships self-directed mistakes and other mistakes related to non-moral skills, call for some form of correction and explanation but not also for the negative affect that typically accompanies these forms of more moral correction.

and able to draw on first-person experience and self-disclosure in relating to a patient: *'You should be congruent, do as you say and say as you do'*; *'a patient will notice if you hide how you feel'*. If one is not congruent, this will often be interpreted by the patient as (and often is) a withdrawal or detachment by the clinician, that is, a severing of the relationship. Congruence comes with involvement and a consequential openness to nurturing reproach in response to perceived off-limits challenging behaviour, and at the same time provides a reason that speaks against withholding and suppressing negative affect. Our interviewees thought that even when they were on the verge of losing control over their emotions in response to extreme and persistent objectionable conduct and should step out of the situation, it was still good practice to explicitly mark their anger and frustration, explain why they had to leave, leave in order to cool down, only then to come back and restore contact again.

2. Sense of safety

Nurturing reproach also helps to provide the service user with a sense of safety, partly because being transparent renders a clinician more predictable and reliable in the eyes of the service user. Being congruent is instrumental to the service user's feeling of safety. Unsuccessful attempts to hide one's negative feelings will create more rather than less confusion and anxiety for the service user about social boundaries and renders the clinician less reliable and predictable (e.g. incongruence between one's verbal message and one's facial expression, bodily composure and tone of voice). But a sense of safety is also provided because by labelling certain behaviours as unacceptable, the service user is presented with boundaries to what they can do. Our interviewees considered nurturing reproach to be crucial for such boundary setting. They made comments such as, *'They may notice that I am a bit angry when they really cross a line; if you say something too calmly they may think... is this actually a problem? Is this ok or not?'* Nurturing reproach illustrates that the clinician as a person has and respects clear socio-moral boundaries. This provides a holding environment that reduces anxiety; the patient is included in (rather than excluded from) a shared practice and shown what the rules of this practice are at the same time.

3. Fostering agency

In the interviews clinicians refer to the different ways in which an expression of anger may encourage a service user's self-governing approach towards objectionable patterns of behaviour (Brandenburg, 2017). To illustrate, one clinician remarked, *'My experience tells me many patients are pampered, and that this makes them in fact more insecure and unsure of themselves ... if you don't correct someone, you, without maybe meaning to do so, will give them the message that they cannot do this by themselves'* and *'You should place people in a position from which they can learn to improve.'* One

thereby also helps to avoid victimization and disempowerment. Another clinician said in relation to this, *'It is your therapeutic responsibility to recover the person's autonomy, make them feel human more than victim or underdog.'* Calling upon the person to have a self-governing approach does not imply, of course, that the service user will not be offered the help and support needed. What is relevant is that correcting a person may help to encourage the person's proactive attitude in this process of supported recovery.

A related manner, in which nurturing reproach may foster agential capacities, is by inspiring a process of reflection on one's conduct. One clinician put this succinctly by saying that after he expressed to be angered by objectionable conduct, *'I expect that responsibility is then placed on the stage, that this expression initiates internal dialogue.'*²³

4. Recognition

A number of clinicians mentioned how, by showing that someone affects them –rather than concealing this – they *'take a person seriously, by showing some things simply can't be done'* or see *'the person, in some sense, as an equal and valuable community member'* or simply *'express [that] you respect the person.'* This notion of recognition may be elusive if not intuitively grasped by the reader. There is a sense in which nurturing reproach conveys the message that the service user is recognized and treated as a member a shared norm-guided practice. One is an equal member of this practice in the sense that one can be called upon to recognize and work towards abiding by interpersonal norms, and is someone who can call upon other members to do the same (e.g. Brandenburg 2019).

5. Exemplarity

Lastly, controlled expressions of nurturing reproach are also considered to provide instructive examples. First, they serve as a model of how people outside the clinic may react to challenging behaviour. This is something that service users ought to know and be able to respond to when they leave the clinical setting. A concern voiced by our interviewees was that service users become more institutionalized when not directly confronted with a person's corrective response to their challenging behaviour. One interviewee said, *'At times you only realize how their behaviour made you feel on your way home. I can understand a clinician's choice to defer a response... But still I worry; it makes me think, outside of the clinic you won't get deferred responses and to what extent is this then instructive?'*

²³ This reason connects to an important insight central to feminist ethics: human autonomy is relational because good social relationships support one's ability to live one's life in the light of carefully considered and endorsed values, goals and plans.

Second, nurturing reproach also serves as an illustration of how service users themselves may learn to experience and express nurturing reproach, directed at self or others in healthier and more adaptive ways than they are used to. Many service users associate the negative affects in corrective attitudes with being bad, deserving punishment and being excluded. Nurturing reproach can provide a corrective experience of being held accountable for misbehaviour while still being accepted and worthy of consideration and empathy. One clinician remarked that deferred suppressed responses would *'fail to exemplify how it may be perfectly acceptable for a patients to address their feelings and talk and think about them'*. He added that in his experience many patients are surprised by the realisation that it is possible and acceptable to address and talk about negative emotions, which further stresses the need for showing how one may experience and express nurturing reproach. By setting an example, clinicians show service users when and how negative emotions can be expressed in a safe, non-disruptive way in social relationships. Internalization of these attitudes may furthermore enable service users to treat themselves in a more compassionate and constructive way in response to their own failures to meet moral norms or standards.

7.4 Comparing nurturing reproach and detached blame

In this section, we compare nurturing reproach to Pickard's notion of detached blame from a clinical stance. We firstly discuss some similarities between the two accounts. Secondly we explore how elements of our account could possibly supplement the clinical stance. Thirdly we address a crucial difference between the two accounts.

Both nurturing reproach and detached blame are a response to objectionable conduct that differs from condoning or 'rescuing'. As we discussed above, the clinicians interviewed believed that one important reason that speaks in favour of correcting a person is that it fosters agency. A more 'rescuing' or condoning approach fails to facilitate such empowerment and moral uptake. Nurturing reproach is thus similar to detached blame in that it avoids rescuing and condoning in order to maintain and develop the patient's agency.

Does nurturing reproach also distinguish itself from affective blame? The clinicians we interviewed did report to experience and express certain negative affect in response to objectionable conduct. They also believe that there is a place for mild forms of anger within a clinical setting. However, these negative attitudes in response to norm-transgressions do not amount to blaming according to them. One clinician remarked that *'Blame is not the right term ... People often acquire these problematic habits because nobody ever corrects them.'* *What I do is indicate my limits; express that this is something I will not tolerate.'* She considered her affective responses to have a

‘corrective’ character rather than a retributive or blaming character. *‘I believe that you should always correct someone, just as one would do for a child. If a two year old starts to throw things through the room you would also correct them. This is how I see my responses to challenging conduct, and in such cases you are not blaming someone, I would not call that blame.’*

This denunciation of blaming seems consistent with Pickard’s account. According to Pickard affective blame is a specific type of negative affect characterized by a sense of entitlement. These feelings are inconsistent with sympathy and compassion and, therefore, do not belong in a therapeutic relationship. The clinicians we interviewed also rejected those particular emotional responses because they believed affective responses should not undermine therapeutic recovery and should be consistent with their professional role.

Other types of negative affect can however be consistent with sympathy and compassion. There is an important difference between the nurturing and educative manners in which one would affectively correct people that are in one’s care, and the resentment or indignation one may feel towards say a friend, colleague or partner who has wronged you or someone close to you. The second type is considered incompatible with a therapeutic relationship in a context of mental health care, but the first is not.

It is worth noting here that some philosophical accounts of blame seem to include *any type* of nurturing negative response that is triggered by (perceived) objectionable conduct towards others (Fricker, 2016; McGeer, n.d.). This raises the question whether these definitions of blame are inconsistent with natural language definitions, or whether the definition of blame within this particular practice are different from some philosophical accounts of blame and maybe other practices. This is not something we can pursue in this chapter. Our account suggests that mild negative responses to objectionable conduct are acceptable and sometimes even desirable in contexts of mental health care. The clinicians we interviewed would however *not* describe these responses as blaming.

Pickard’s notion of *detached* blame suggests no involvement of negative affect at all and alludes to a cool and collected response to objectionable conduct. Our account suggests that there are dangers to such a response if it would involve repressing and obfuscating any negative affect. Such a response may then, at least for this service user group, be incongruent and lack the sense of safety that comes with a more transparent form of relating to the person. We do not mean to deny here that it is possible to have a *genuine* calm and collected response to a person. Characterizing a detached response as the guideline may have undesirable consequences in all those cases where a clinician *is* negatively affected by objectionable conduct and the service user

would be sensitive to affective appeal. And such cases, according to our interviewees, are common.

On a charitable reading of Pickard, however, mild negative affect is not necessarily inconsistent with her notion of responsibility without blame. Although Pickard does not explicitly address these milder negative affective responses they may be consistent with her overall account because these responses would not qualify as *affective blame* on Pickard's definition, and it is only affective blame that she explicitly considers incompatible with the therapeutic relationship. Like the clinical stance, nurturing reproach distinguishes itself from Pickard's definition of affective blame. There is no justification for a hostile form of anger that comes with the feeling of the other person being deserving of this anger and the clinician being entitled to this anger. Nurturing reproach then helps to demarcate this type of blame from other affective responses that could be acceptable and maybe even desirable in the context of a psychiatric clinic. Nurturing reproach and detached blame are thus similar in so far as both are responses to objectionable conduct that find a mean between (some definitions of) affective blame on the one hand and condoning or rescuing on the other hand.

There are other ways in which our account can supplement Pickard's clinical stance. The reasons discussed in the last section provide additional considerations that justify forms of responsibility without blame. First they provide some additional ways in which forms of responsibility without blame may be conducive of therapeutic recovery. Pickard explains why the clinical stance is conducive of therapeutic recovery by contrasting it to two other types of responses that undermine treatment: rescuing and affective blaming. A rescuing response obfuscates the possibility that the service user can take control and engage in her own recovery process. A blaming response on the other hand fails to empower the service user because it tends to make the person feel rejected and incurs a vicious cycle of guilt. Both undermine the service user's self-governing approach to recovery as well. The clinical stance is conducive to recovery *because* it fosters or facilitates agency in recovery.

Our account supplements this justification by providing more considerations that explain how responsibility without blame can be conducive to therapeutic aims. Fostering agency in recovery is among those considerations. But in addition congruence, a sense of safety, recognition and exemplarity may also explain how responsibility without blame can secure therapeutic success (see previous section).

Furthermore our justificatory framework may supplement the clinical stance because over and above facilitating recovery there are *intrinsic* reasons that speak in favour of a response

that mediates between rescue and blame (see ‘recognition’ in the previous section). The interviewed clinicians do not only engage in nurturing reproach *in order* to achieve certain therapeutic purposes. Nurturing reproach is also considered a sincere response to a service user’s behaviour that embodies respect for the person. It is a respectful form of relating because it is congruent towards the service user and recognizes him/her as a member of a shared norm-guided community. Forms of holding responsible without blame are then appropriate responses, not *only* because they contribute to therapeutic recovery but also because they amount to a respectful form of relating to the person.

There are similarities between nurturing reproach and Pickard’s notion of detached blame from a clinical stance, and our account can supplement the clinical stance, but there is also a crucial difference between the two accounts. On our account, *detached blame (an attribution of blameworthiness) neither explains nor justifies practices of holding responsible* that are embodied in attitudes of nurturing reproach. Recall that on Pickard’s account the notion of detached blame serves to explain *how* we can hold a person responsible without also ‘blaming’ the person: we may judge the person to be blameworthy without getting emotionally exercised about this fact (2013, pp. 1142–1146). On Pickard’s account the clinical stance is also justified to the extent that the person is responsible for this conduct and in a *detached* sense blameworthy for it too.

As we explained in the previous section, the clinicians we asked to explain and justify their responses of nurturing reproach to objectionable conduct by a service user, generally did not believe this person to be blameworthy or to have the level of understanding and control that may be considered sufficient for blameworthiness. To provide one example, a senior nurse reported that a week earlier a female service user on her ward had turned down the opportunity to speak to two of her colleagues in order to arrange a meeting with her. The woman refused to listen or talk to these two colleagues, and her behaviour towards them became more and more verbally abusive. The senior nurse was at that point taking a break in a common room. She overheard her interacting with her colleagues in the hallway and considered her treatment of her colleagues out of line. The interviewee said, ‘*I was totally done with this behaviour. It crosses a line. You cannot treat other people this way and you cannot get things done by me that way either. I think it is appropriate to show this and let someone know.*’ She took the woman aside and told her, ‘*You are here because you want to receive treatment, but you refuse to speak to two of my colleagues. Do you want to cooperate and work on your treatment? It is up to you. But that requires you treat them with more respect.*’ She then left the situation and let her colleagues take over.

This clinician clearly affectively corrected this service user for verbally abusing other staff members, but when asked whether she thought this person could be expected to have behaved

better under the circumstances, she replied in the negative. Elaborating on her answer, she pointed out that this person's social-emotional intelligence was disproportionately low in relation to her above-average IQ (not uncommon in cases of autism spectrum disorder). Seen through the lens of poor social-emotional development, this kind of regressive behaviour is understandable and even to be expected in stressful situations. She added that this service user had been repeatedly hospitalized elsewhere where such behaviour was reinforced. According to the interviewee, this service user could not be considered blameworthy for her conduct given those agency-compromising factors. Yet nurturing reproach was considered desirable for the reasons discussed above.

On our account, (detached) blame (worthiness) neither justifies nor explains the responses of the interviewed clinicians towards service users' objectionable conduct. Rather, the reasons that were listed in the previous section (congruence, sense of safety, fostering agency, recognition and exemplification) together seem to provide sufficient justification for why a service user should be subject to nurturing reproach when she transgresses an important moral norm. But if nurturing reproach is a form of holding the person responsible, how can it be appropriate to do so if the person is not blameworthy for her objectionable conduct?

Recall that Pickard introduces to concept of detached blame also in order to explain how responsibility without blame is possible. The interviewed clinicians in our study considered the notion of (detached) blame to mischaracterize their corrective responses to objectionable conduct in this clinical setting. They either did not know whether the degree of capacities for choice and control implied by Pickard's discussion of detached blame was applicable, or they were positive that it did not apply in the particular situation of the service user. In light of these considerations, the clinicians interviewed often did not attribute to the service users in question blameworthiness for their objectionable conduct. The question then is how we can explain 'responsibility without blame' embodied in nurturing reproach, without making use of some attenuated notion of blame in relation to the objectionable conduct under consideration.

We take it that nurturing reproach does not imply that one considers the person to be responsible (and blameworthy) *for* her objectionable behaviour. When one engages in nurturing reproach, one allocates other types of agential abilities and responsibilities to the person that are relevant *because of* the objectionable behaviour. For example, the agency and responsibility to continue to work on one's own recovery, and the communicative skills and responsibilities to receive and respond to another person's attitude of nurturing reproach. These abilities and responsibilities do *not* amount to the person *being responsible for her objectionable conduct*. They instead

refer to a responsibility that service users have to work together with their clinician towards recovery and to the responsibilities they have to motivate and enable themselves and others to display due regard for one another within a given community. The clinician appeals to the service user and the latter is thereby, in a sense, 'held responsible' by the clinician. However, this form of holding responsible does not amount to detached blame towards the objectionable conduct; the patient is held responsible for other more future directed concerns.

It is of course conceivable that the patients in the clinic are sometimes blameworthy for what they do. But the clinicians we interviewed preferred not to attribute blameworthiness unless they could be quite certain about a person's level of understanding and control over her own objectionable conduct. The service users' inability to control for certain forms of objectionable behaviour was often part of the reason why they sought treatment in this clinical setting. Contrary to the outside world, the clinic is then a place where the person is not *yet* considered to have the responsibility to control such behaviour but is helped to become able to do so.

How does this future directed form of holding responsible expressed by nurturing reproach relate to Pickard's notion of detached blame? It remains to be seen whether our account of nurturing reproach and Pickard's understanding of detached blame from the clinical stance are two competing approaches, or rather highlight different aspects of an overarching therapeutic stance. One possibility is that the responses considered appropriate depend on the type of disorder and the particular context of intervention. Perhaps detached blame is generally more fitting for objectionable behaviour originating in the interpersonal dynamics of (borderline) PD, whereas nurturing reproach might be more appropriate for challenging behaviour against the background of (certain forms of) ASD. For example, it could be theorized that, due to the nature of borderline personality disorder and the specific interpersonal hypersensitivities this encompasses, detached blame is more appropriate and conducive to recovery than forms of nurturing reproach. More likely, however, the appropriateness of each kind of response is not determined by (DSM) diagnosis, but by individual characteristics such as the specific socio-emotional, cognitive and interpersonal attachment profile of the individual service user, as well as by the specific context of the objectionable conduct and the clinician's response.²⁴

If we are right, there is place for both detached blame and nurturing reproach in different therapeutic situations or settings. We hypothesise that detached blame is the more appropriate

²⁴ In this context it should be noted that research suggests that there is significant co-morbidity of (borderline) PD and ASD (based on DSM IV/5 criteria) and also symptomatic overlap (e.g. Hofvander et al., 2009; Lugnegård, Hallerbäck, & Gillberg, 2012; Ryden, Ryden, & Hetta, 2008). Recent research shows relatively high scores on ASD-traits in people diagnosed with BPD, also suggesting overlap of the two diagnostic constructs (Dudas et al., 2017).

response in reaction to conduct that is explained by a person's attitude towards recovery and attitude towards the therapeutic relationship as part of recovery. Pearce and Pickard describe how a decision or choice on the part of the patient is needed in order for the patient to recover; they refer to this as 'the will to recover'(Pearce & Pickard, 2010b). When objectionable conduct is primarily explained by the person's lack of commitment to recovery and lack of acceptance of the therapeutic relationship as part of recovery, the service user is less likely to be sensitive to the appeal embodied in nurturing reproach. Rather what is needed is for the patient to decide or choose to work towards recovery, and, relatedly, to accept and trust the therapeutic relationship and the value of a collaborative effort in the therapeutic process towards recovery. This decision or choice on the part of the service-user may be absent for good reasons. As Pickard points out, service users may have had histories of trauma and abuse, which may explain why they are sceptical about the possibility of recovery, and unwilling to trust or accept someone who offers them help. Lack of the decision to recover may be explanatory of objectionable conduct itself, and it may be explanatory of a refusal to respond to a clinician's nurturing reproach.

We take it a detached response that stresses responsibility is more appropriate in these cases because one thereby recognizes the person's choice and control in this matter and allows the person to take this position and reject the therapeutic relationship. One furthermore remains open to therapeutically relating to the person and accepting of the person, without forcing this relationship on the person. We agree with Pickard that the person should at least be seen as *responsible* for refusing to partake in recovery and to accept the therapeutic relationship as a part of this trajectory to recovery.²⁵ This is a choice the service-user makes and something that only she can decide to change and would have to change for therapy to become effective. The hypothesis is then that a detached response that attributes responsibility without thereby rejecting the person, or denying the possibility of a therapeutic relationship, would be a desirable response for service-users with this profile.

Nurturing reproach may be a less desirable and less effective response when the therapeutic relationship is under pressure in this way. Firstly, if a person mistrusts or challenges the therapeutic relationship and is unwilling to engage in a therapeutic process, nurturing reproach is unlikely to be effective. A service user is only open to the clinician's evaluation and disapproval of her conduct if she considers herself to stand in a particular relationship to the

²⁵ One may of course still wonder if the person should be considered blameworthy for taking this attitude towards therapy and towards clinicians. If in this service-user's past, other people have proven to be untrustworthy and abusive, there arguably have good reason to be mistrustful of any other person. We suspect that sometimes this refusal is understandable and would not merit an attribution of blameworthiness for such conduct over and above responsibility for such conduct, but do not have the space to discuss this issue more in depth.

clinician e.g. as together participating in a therapeutic process that is aimed at her recovery. Absent this underlying acceptance of the clinician as someone who can be trusted to be willing and able to help you, one will probably not take this clinician's appeal to be acceptable and helpful. Secondly, nurturing reproach may be considered unduly paternalistic in these contexts, because it aims to engage and improve the person in a shared therapeutic process even though the person herself rejects this participation. Nurturing reproach aims to 'develop' the person's abilities by engaging a person in a therapeutic process. But if the person is, at this point in time, unwilling to participate in this process, nurturing reproach would seem to go against the person's own will. As such it would amount to a paternalistic response that fails to respect the person's autonomy and responsibility. The person would be better respected by an acknowledgment of the disagreement regarding therapy or the therapeutic relation.

In the contrary situation, however, we believe the opposite is true. Detached blame is less respectful and conducive to therapy when a service user is committed to recovery, involved in the therapeutic process, and accepts the role of the clinician in this process, the objectionable conduct notwithstanding. In those cases nurturing reproach stresses that involvement and helps the person to take, or keep taking, the future directed responsibilities that enable the person to work on their recovery.

If the above hypothesis is correct, the two different therapeutic responses are compatible and both have value in mental health care contexts. Further research is required to test this hypothesis and to evaluate these two types of 'responsibility without blame'. Other variables may have to be taken into account. Different socio-cultural contexts may place different value on some of the considerations we discussed and/or may have different concepts of therapeutic recovery. The specific meaning of emotional expression may also subtly vary, depending on (sub)culture and possibly patient-group. Our account needs to be tested from more than one perspective. Importantly, the perspective of the service user is missing in the present study. In further support of therapeutic responses to objectionable behaviour, the perspective of the service users involved is indispensable. Despite this need for further research, we hope to have provided a critical comparison and the beginnings of a research project into therapeutic forms of holding responsible in mental health care contexts.

7.5 Conclusion

In this chapter we provided an account of therapeutic responses of clinicians to objectionable conduct by service users in a clinical setting for adults with ASD. Based on an explorative qualitative study among clinicians who work in this setting, we developed an understanding of

these responses in terms of nurturing reproach. We compared the results of this study to Pickard's notion of 'detached blame' in her clinical stance account. Although there are similarities and ways in which our account may complement Pickard's clinical stance, there is also an important difference. Pickard's notion of detached blame is supposed to explain how one can hold someone responsible for objectionable behaviour without affectively blaming the person for this behaviour. In contrast, our understanding of nurturing reproach embodies controlled expression of negative affect on the part of the clinician that appeals to a set of future-directed responsibilities of the service user, but without attributing any kind of responsibility or (detached) blame to the person regarding the objectionable conduct to which the correction was a response.

As we pointed out in the previous section, it remains to be seen whether our account of nurturing reproach and Pickard's understanding of detached blame from the clinical stance are incompatible, or rather highlight different aspects of an overarching therapeutic stance. We hypothesised an important contextual factor to take into account is to what extent the service user trusts and accepts the therapeutic relation in the process of recovery, as this suggest the person can constructively respond to the appeal on future responsibilities in therapy that is embodied in nurturing reproach. If the service user at that particular moment is not receptive to this appeal, it seems that nurturing reproach is the wrong kind of response. In situations where service users for example challenge or do not trust the therapeutic relationship, affectively charged corrections in response to objectionable conduct may be counterproductive, and detached blame may be the more appropriate response.

Our proposed account critically compares to Pickard's clinical stance and provides the beginnings of a more extensive research project on therapeutic forms of holding responsible in mental health care. We should keep in mind that our own and Pickard's findings are based on what is perceived as best practice in particular clinical settings by clinicians working there. The findings should not be extrapolated to other health care settings without question. This explorative account needs further research from more than one perspective, in particular the perspective of the service user.

8 Conclusion

8.1 Summarising conclusion

In this thesis I conceptualised a form of holding responsible that I refer to as a ‘nurturing stance’. From a nurturing stance we do not blame and we do not attribute blameworthiness to the person who has done harm, because we do not consider the person is (yet) independently able to respond to the relevant moral reasons. Our interactions that are expressive of this stance are instead explained by the fact that we consider a person able to communicate about objectionable conduct and able to accept future-directed responsibilities to develop or manage their abilities, in response to the harm that was done. The stance thereby recognises a person’s abilities to engage in a moral conversation and their abilities to be involved in a process of moral development or self-management.

The philosophical literature on responsibility has neglected this sphere of our human interactions. There is a strong focus on how we may hold individuals responsible when they are already independently responsive to moral reasons. But not much attention has been paid to what happens when a person is insufficiently competent to be considered blameworthy for what they have done. I discussed the theoretical implications of my account for Strawsonian theories of responsibility in general, certain functional accounts of blame, Shoemaker’s tripartite theory, and Hannah Pickard’s clinical stance. I illustrated how all of these accounts are at least ambiguous and at worst misleading where they concern normative interactions with individuals who are not blameworthy because they are not yet able to by themselves do better under similar circumstances. The nurturing stance provides an approach that resolves these ambiguities and corrects for these misrepresentations. It also renders the literature more inclusive and respectful of individuals whose moral capacities are compromised or underdeveloped.

I discussed two institutional contexts within which my account may be of practical use. First, within the context of criminal law, the nurturing stance provides a way to further extend recognition of the capacities of persons with mental disabilities without unfairly blaming them. Second, within the context of psychiatric care, the nurturing stance shows how a clinician may respect and strengthen a service user’s responsibility when this clinician does not consider the patient to be blameworthy for objectionable conduct on the basis of insufficient capacity.

8.2 Directions for further research

I have a number of relevant directions for further research. Implementing a nurturing stance within the context of criminal law or adopting it within psychiatric therapy would require more research. This research may concern figuring out whether and how implementation is feasible, but also when and to what extent it is morally desirable. When are patients and service users benefitted most by being engaged in a nurturing dialogue, and what are the concerns and interests of other parties when it comes to such implementation? More research can also be done regarding the demarcations of the nurturing stance. The contours of when a nurturing stance is appropriate have been discussed in this thesis. The capacity conditions for being appropriately subjected to a nurturing stance are having the abilities to receive and respond to moral address and to participate in a future-directed process that helps to prevent similar harm from occurring. But one may ask when a person would be minimally sufficiently competent to receive and respond to moral address. This question connects to debates about the moral development of children, animal cognition, and metaethical debates on moral reasoning more generally. It would be interesting to see how these debates impact and are impacted by human interactions that express a nurturing stance.

At the other end of the spectrum of nurturing responses, one may wonder when a person should be considered *too* competent to be subjected to a nurturing stance. Simply put, a nurturing stance is problematic if one believes the person's moral skills are underdeveloped or compromised *without* having good reason to do so. For example, a person's physical disability would be a bad reason for taking a nurturing stance towards the person. Arguably, taking a nurturing stance because a person has been diagnosed with a mental disorder, without knowing if, and if so how, this disorder may affect a person's moral skills, would also be discriminatory. Furthermore, neurotypical individuals may be biased about the moral import of certain reasons. A neurotypical may therefore mistakenly take a nurturing stance towards a person, when in fact the other person *is* able to respond to the relevant reasons, but did not consider those reasons (sufficiently) morally relevant under the circumstances. This mistake need not be harmful because the nurturing stance allows for responses that voice disagreement and open up discussion about the moral import of certain reasons. But, further research can be done on when a nurturing stance would be discriminatory or paternalistic.

So far, the suggested directions for further research concern further elaborations on the work that has been provided in this thesis. There are also new directions for research that could be relevant. I suspect other debates about responsibility may have equally overlooked or

misrepresented the ways in which individuals may be held responsible *because of* harm done, when they lack sufficient capacity to be considered blameworthy *for* the harm that is done. Among them are debates on shared responsibility, and debates on moral luck and moral injury.

First, further research may show how the nurturing stance can contribute to debates on shared responsibility. One of the main issues within the debate on shared responsibility is explaining how – if at all – a group of agents can be blameworthy for harm. I suspect that also in this debate a focus on blame and blameworthiness has prevented philosophers from carefully considering how we can and should respond to certain groups of agents that lacked sufficient abilities for being blameworthy. By engaging a group in a dialogue about the harm that is done, and by appealing to their abilities to take future responsibilities because of this harm, we may help establish capacities that can render a group independently accountable in the future. It would therefore be interesting to inquire how groups of agents can and should sometimes be subjects of a nurturing stance.

Second, it would be interesting to inquire how the nurturing stance may be of use when it comes to addressing issues related to moral luck and moral injury. The problem with moral luck cases is that it seems hard to explain why we would hold it against a person when she does not take any responsibility for serious harm that was done *through no fault of her own*. A famous example is that of a truck driver who hits a child who ran onto the road (Williams, 1982, p. 43). The truck driver in this scenario could not have done anything to avoid hitting the child. One of the questions raised by this scenario is why do we think the truck driver *should* feel sorry if he is not blameworthy for what he has done?

In certain cases of moral injury the opposite may occur. In those cases, a person has fulfilled his duties and has done what was morally required of him, but this has resulted in grave harm being done to others. This may commonly happen in contexts of war. Imagine a case in which a soldier cannot help a mother whose child needs immediate medical assistance because he has to follow certain orders that require him to leave the village now. When leaving the mother and child, the soldier has injured his ‘moral conscience’ and this continues to cause feelings of shame and guilt. One important question to be answered in this debate is why do acts of atonement *help* address the injury, and, relatedly, why does it typically *not help* to tell a person they did nothing wrong?

In cases of moral luck and (certain cases of) moral injury, neither blaming nor excusing seems appropriate. I suspect this may be less puzzling when it is recognized how a person can be responsible because of harm without also being responsible for harm. The nurturing stance is a form of holding responsible that finds a mean between blaming and excusing. It recognizes a

person as someone who can be involved in establishing and living up to future responsibility that arise because of the harm that was done.

It is possible that the nurturing stance, or an adaptation of this stance, can help explain why and how the truck driver should ‘feel sorry’ and, in some way, responsible. One could say that, because of the harm, the truck driver incurs a number of future directed responsibilities that involve: opening a dialogue with the victims, looking into possible future responsibilities that (somewhat) help address the harm that was done, and establishing how he himself can be involved in taking on (some of) those future responsibilities. None of this requires we consider the truck driver blameworthy for what he has done.

Similarly, an (adapted) nurturing stance may help the soldier come to terms with his moral injury in ways that do not excuse him from, or blame him for, the harm that was done. By engaging the soldier in a conversation about how the past harm can be (partly) addressed through him -and others involved- taking certain responsibilities that help prevent similar harm from occurring in the future, we may have found a way to help heal the soldier’s moral injury without unfairly blaming him.

The above suggestions are, of course, preliminary. All I hope to have shown here is that the nurturing stance amounts to a form of holding responsible that may have been overlooked, not just in the debates addressed in this thesis, but also in other responsibility-related debates. This account has the potential to resolve ambiguities in those other debates too.

References

- Andrews, K. (2012). *Do Apes Read Minds?: Toward a New Folk Psychology*. The MIT Press.
- Antony, L. M. (2016). Bias: Friend or Foe? In M. Brownstein & J. Saul (Eds.), *Implicit Bias and Philosophy, Volume 1* (pp. 157–190). Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780198713241.003.0007>
- Australian Neurolaw Database. (2015). [<https://neurolaw.edu.au/cases/10403>]. Retrieved 27 February 2019, from
- Baron-Cohen, S. (1997). *Mindblindness: An Essay on Autism and Theory of Mind* (Revised ed. edition). Cambridge, Mass.: A Bradford Book.
- Bennett, C. (2008, August). The Apology Ritual by Christopher Bennett.
<https://doi.org/10.1017/CBO9780511487477>
- Bermond, B., Clayton, K., Liberoval, A., Luminet, O., Maruszewski, T., Bitti, P. E. R., ... Wicherts, J. (2007). A cognitive and an affective dimension of alexithymia in six languages and seven populations. *Cognition and Emotion*, 21(5), 1125–1136.
<https://doi.org/10.1080/02699930601056989>
- Bird, G., & Viding, E. (2014). The self to other model of empathy: providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. *Neuroscience and Biobehavioral Reviews*, 47, 520–532.
<https://doi.org/10.1016/j.neubiorev.2014.09.021>
- Bloom, P. (2012). Moral nativism and moral psychology. In Mario Mikulincer and Philip R. Shaver (Ed.), *The social psychology of morality: Exploring the causes of good and evil* (pp. 71–89).
- Brandenburg, D. (unpublished). Interview Data. *Available upon Request*.
- Brandenburg, D. (2016). Implicit attitudes and the social capacity for free will. *Philosophical Psychology*, 29(8), 1215–1228. <https://doi.org/10.1080/09515089.2016.1235263>

- Brandenburg, D. (2017). The Nurturing Stance: Making Sense of Responsibility without Blame. *Pacific Philosophical Quarterly*, 99(S1). <https://doi.org/10.1111/papq.12210>
- Brandenburg, D. (2019). Inadequate Agency and Appropriate Anger. *Ethical Theory and Moral Practice*. <https://doi.org/10.1007/s10677-019-09982-w>
- Brink, D. O., & Nelkin, D. K. (2013). *Fairness and the Architecture of Responsibility* (SSRN Scholarly Paper No. ID 2313826). Rochester, NY: Social Science Research Network.
- Calhoun, C. (1989). Responsibility and Reproach. *Ethics*, 99(2), 389–406.
- Clarke-Stewart, K. A. (1973). Interactions between Mothers and Their Young Children: Characteristics and Consequences. *Monographs of the Society for Research in Child Development*, 38(6/7), 1–109. <https://doi.org/10.2307/1165928>
- Coates, D. J., & Tognazzini, N. A. (2012). *The Contours of Blame*. Oxford University Press.
- Coates, D. J., & Tognazzini, N. A. (2013). The Contours of Blame. In D. J. Coates & N. A. Tognazzini (Eds.), *Blame: Its Nature and Norms* (pp. 3–26). Oxford University Press.
- Committee on the Rights of Persons with Disabilities . United Nations. Article 12 - Equal recognition before the law (2014). Retrieved from <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities/article-12-equal-recognition-before-the-law.html>
- Craigie, J. (2015). Against a singular understanding of legal capacity: Criminal responsibility and the Convention on the Rights of Persons with Disabilities. *International Journal of Law and Psychiatry*, 40, 6–14. <https://doi.org/10.1016/j.ijlp.2015.04.002>
- D’Arms, J., & Jacobson, D. (2000). The Moralistic Fallacy: On the ‘Appropriateness’ of Emotions. *Philosophy and Phenomenological Research*, 61(1), 65–90. <https://doi.org/10.2307/2653403>
- Deigh, J. (2011). Reactive Attitudes Revisited. In C. Bagnoli (Ed.), *Morality and the Emotions*.

- Dudas, R. B., Lovejoy, C., Cassidy, S., Allison, C., Smith, P., & Baron-Cohen, S. (2017). The overlap between autistic spectrum conditions and borderline personality disorder. *PLOS ONE*, 12(9), e0184447. <https://doi.org/10.1371/journal.pone.0184447>
- Duff, A. (2002). Punishing the Young. In I. Weijers & A. Duff (Eds.), *Punishing Juveniles: Principle and Critique*. Hart.
- Duff, A. (2003). *Punishment, Communication, and Community*. Oxford, New York: Oxford University Press.
- Duff, R. A. (2011). *Responsibility, Restoration, and Retribution*. Oxford University Press.
- Emerson, E. (2001). *Challenging Behaviour: Analysis and Intervention in People with Severe Intellectual Disabilities* (2 edition). Cambridge ; New York: Cambridge University Press.
- Fiebich, A. (2017). Pluralism, social cognition, and interaction in autism. *Philosophical Psychology*, 30(1–2), 161–184. <https://doi.org/10.1080/09515089.2016.1261394>
- Fiebich, A., & Coltheart, M. (2015). Various ways to understand other minds: Towards a pluralistic approach to the explanation of social understanding. *Mind & Language*, 30(3), 235–258.
- Fischer, J. M., & Ravizza, M. (1998). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge University Press.
- Fischer, J. M., & Ravizza, M. (2000). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge University Press.
- Freckelton, I. R., & List, D. (2009). *Asperger's Disorder, Criminal Responsibility and Criminal Culpability* (SSRN Scholarly Paper No. ID 1810117). Rochester, NY: Social Science Research Network.
- Fricker, M. (2016). What's the Point of Blame? A Paradigm Based Explanation. *Nous*, 50(1), 165–183. <https://doi.org/10.1111/nous.12067>
- Fricker, M. (2018). Ambivalence About Forgiveness. *Royal Institute of Philosophy Supplement*, 84, 161–185.

- Gopnik, A. (2009). *The Philosophical Baby: What Children's Minds Tell Us About Truth, Love, and the Meaning of Life* (1st edition). New York: Farrar, Straus and Giroux.
- Grandin, D. T., Barron, S., & Barron, S. (2005). *Unwritten rules of social relationships : decoding social mysteries through the unique perspectives of autism*. Arlington, TX: Future Horizons.
- Greene, R. (2014). *Explosive Child, The: A New Approach For Understanding And Parenting Easily Frustrated, Chronically Inflexible Children* (Revised, Updated edition). New York: HarperPb.
- Grusec, J. E., & Goodnow, J. J. (1994). Impact of parental discipline methods on the child's internalization of values: A reconceptualization of current points of view. *Developmental Psychology*, 30(1), 4–19. <https://doi.org/10.1037/0012-1649.30.1.4>
- Grusec J.E., T Dix, R Mills. (1982). The effects of type, severity and victim of children's transgressions on maternal discipline. *Canadian Journal of Behavioral Science*, 14, 276–289.
- Hayes, S. (1994). The criminal law and the person with intellectual disability. *Australia and New Zealand Journal of Developmental Disabilities*, 19(4), 287–292. <https://doi.org/10.1080/07263869400035341>
- Helm, B. W. (2012). Accountability and Some Social Dimensions of Human Agency. *Philosophical Issues*, 22(1), 217–232. <https://doi.org/10.1111/j.1533-6077.2012.00226.x>
- Hieronymi, P. (2004). The Force and Fairness of Blame. *Philosophical Perspectives*, 18(1), 115–148.
- Hirsch, A. von. (1996). *Censure and Sanctions*. Oxford University Press.
- Hoffman, M.L. (1983). Affective and cognitive processes in moral internalization: An information processing approach. In E.T. Higgins, D. Ruble, & W. Hartup (Ed.), *Social Cognition and Social Development: A Socio-Cultural Perspective*. (pp. 236–274). New York: Cambridge University Press.
- Hofvander, B., Delorme, R., Chaste, P., Nydén, A., Wentz, E., Ståhlberg, O., ... Leboyer, M. (2009). Psychiatric and psychosocial problems in adults with normal-intelligence autism spectrum disorders. *BMC Psychiatry*, 9, 35. <https://doi.org/10.1186/1471-244X-9-35>

- Huebner, B. (2016). Implicit Bias, Reinforcement Learning, and Scaffolded Moral Cognition. In M. Brownstein & J. Saul (Eds.), *Implicit Bias and Philosophy, Volume 1* (pp. 47–79). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198713241.003.0003>
- Jaarsma, P., & Welin, S. (2012). Autism as a natural human variation: reflections on the claims of the neurodiversity movement. *Health Care Analysis: HCA: Journal of Health Philosophy and Policy*, 20(1), 20–30. <https://doi.org/10.1007/s10728-011-0169-9>
- Keltner, D., & Haidt, J. (2003). Social Functions of Emotion at Four Levels of Analysis. *Cognition and Emotion*, 13(5), 505–521.
- Kennett, J. (2002). Autism, Empathy and Moral Agency. *The Philosophical Quarterly*, 52(208), 340–357. <https://doi.org/10.1111/1467-9213.00272>
- Kennett, J. (2007). Mental Disorder, Moral Agency, and the Self. In B. Steinbock (Ed.), *The Oxford Handbook of Bioethics* (pp. 90–113). Oxford University Press.
- Kennett, J., Vincent, N. A., & Snoek, A. (2015). Drug Addiction and Criminal Responsibility. In J. Clausen & N. Levy (Eds.), *Handbook of Neuroethics* (pp. 1065–1083). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-4707-4_71
- Killen, M., Smetana, J. G., & Smetana, J. (Eds.). (2005). *Handbook of Moral Development* (1 edition). Psychology Press.
- Klin, A., & Jones, W. (2006). Attributing social and physical meaning to ambiguous visual displays in individuals with higher-functioning autism spectrum disorders. *Brain and Cognition*, 61(1), 40–53. <https://doi.org/10.1016/j.bandc.2005.12.016>
- Lacey, N., & Pickard, H. (2013a). From the Consulting Room to the Court Room? Taking the Clinical Model of Responsibility Without Blame into the Legal Realm. *Oxford Journal of Legal Studies*, 33(1), 1–29.
- Lacey, N., & Pickard, H. (2013b). From the Consulting Room to the Court Room? Taking the Clinical Model of Responsibility Without Blame into the Legal Realm. *Oxford Journal of Legal Studies*, 33(1), 1–29. <https://doi.org/10.1093/ojls/gqs028>

- Lacey, N., & Pickard, H. (2015). To Blame or to Forgive? Reconciling Punishment and Forgiveness in Criminal Justice. *Oxford Journal of Legal Studies*, 35(4), 665–696.
- Levy, N. (2016). Implicit Bias and Moral Responsibility: Probing the Data. *Philosophy and Phenomenological Research*, n/a-n/a. <https://doi.org/10.1111/phpr.12352>
- Lewis, M., Haviland-Jones, J. M., & Barrett, L. F. (Eds.). (2010). *Handbook of Emotions, Third Edition* (3rd edition). The Guilford Press.
- Lugnegård, T., Hallerbäck, M. U., & Gillberg, C. (2012). Personality disorders and autism spectrum disorders: what are the connections? *Comprehensive Psychiatry*, 53(4), 333–340. <https://doi.org/10.1016/j.comppsy.2011.05.014>
- Macnamara, C. (2015a). Reactive Attitudes as Communicative Entities. *Philosophy and Phenomenological Research*, 90(3), 546–569. <https://doi.org/10.1111/phpr.12075>
- Macnamara, C. (2015b). Reactive Attitudes as Communicative Entities. *Philosophy and Phenomenological Research*, 90(3), 546–569. <https://doi.org/10.1111/phpr.12075>
- McGeer, V. (2008). Varieties of moral agency: lessons from autism (and psychopathy). In W. Sinnott-Armstrong (Ed.), *Moral Psychology, The neuroscience of morality: Emotion, disease and development* (ed., Walter Sinnott-Armstrong) (Vol. 3). MIT Press.
- McGeer, V. (2011). Co-reactive attitudes and the making of moral community. In R. Langdon & C. Mackenzie (Eds.), *Emotions, Imagination, and Moral Reasoning* (1 edition). New York, NY: Psychology Press.
- McGeer, V. (2013). Civilizing Blame. In D. J. Coates & N. A. Tognazzini (Eds.), *Blame: Its nature and norms*. Oxford University Press.
- McGeer, V. (2014). P. F. Strawson's Consequentialism. In D. Shoemaker & N. Tognazzini (Eds.), *Oxford Studies in Agency and Responsibility, Volume 2* (pp. 64–92). Oxford University Press.
- McGeer, V. (2018). Scaffolding agency: A proleptic account of the reactive attitudes. *European Journal of Philosophy*, 0(0), 1–23. <https://doi.org/10.1111/ejop.12408>

- McGeer, V. (n.d.). Scaffolding agency: A proleptic account of the reactive attitudes. *European Journal of Philosophy*, 0(0). <https://doi.org/10.1111/ejop.12408>
- McGeer, V., & Pettit, P. (2015). The Hard Problem of Responsibility. In D. Shoemaker (Ed.), *Oxford Studies in Agency and Responsibility* (pp. 160–188). Oxford University Press.
- McKenna, M. (2012). *Conversation and Responsibility*. Oxford University Press.
- McKenna, M. (2017). Power, Social Inequities, and the Conversational Theory of Moral Responsibility. In K. Hutchison, C. Mackenzie, and M. Oshana (Eds.), *Social Dimensions of Moral Responsibility*. Oxford: Oxford University Press.
- McKenna, M., & Kozuch, Benjamin. (2015). Free Will, Moral Responsibility, and Mental Illness. In D. D. Moseley & Gala, Gary (Eds.), *Philosophy and Psychiatry: Problems, Intersections and New Perspectives* (pp. 89–113). Routledge.
- McKenna, M. S. (1998). The Limits of Evil and the Role of Moral Address: A Defense of Strawsonian Compatibilism. *The Journal of Ethics*, 2(2), 123–142. <https://doi.org/10.1023/A:1009754626801>
- Menges, L. (2017). The Emotion Account of Blame. *Philosophical Studies*, 174(1), 257–273.
- Milam, P.-E. (2017). In defense of non-reactive attitudes. *Philosophical Explorations*, 0(0), 1–14. <https://doi.org/10.1080/13869795.2017.1298830>
- Milam, P.-E. (n.d.). Reasons to forgive. *Analysis*. <https://doi.org/10.1093/analys/any017>
- Morriss, P. (2002). *Power: A Philosophical Analysis*. Manchester University Press.
- Murphy, J. G. (1979). *Retribution, Justice, and Therapy: Essays in the Philosophy of Law*. Springer Netherlands. Retrieved from <https://www.springer.com/gp/book/9789027709981>
- Nichols, S. (2007). After Incompatibilism: A Naturalistic Defense of the Reactive Attitudes*. *Philosophical Perspectives*, 21(1), 405–428. <https://doi.org/10.1111/j.1520-8583.2007.00131.x>
- Nichols, S., Kumar, S., Lopez, T., Ayars, A., & Chan, H.-Y. (2016). Rational Learners and Moral Rules. *Mind & Language*, 31(5), 530–554. <https://doi.org/10.1111/mila.12119>

- Nichols, S., & Stich, S. P. (2003). *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. Oxford, New York: Oxford University Press.
- Osher, D., Bear, G. G., Sprague, J. R., & Doyle, W. (2010). How can we improve school discipline? *Educational Researcher*, 39(1), 48–58.
- Parish v DPP, No. VSC 494 (Supreme court of victoria 2007). Retrieved from <https://jade.io/j/?a=outline&id=18096>
- Pearce, S., & Pickard, H. (2010a). Finding the Will to Recover: Philosophical Perspectives on Agency and the Sick Role. *Journal of Medical Ethics*, 36(12), 831–833.
- Pearce, S., & Pickard, H. (2010b). Finding the will to recover: philosophical perspectives on agency and the sick role. *Journal of Medical Ethics*, 36(12), 831–833. <https://doi.org/10.1136/jme.2010.035865>
- Pereboom, D. (2009). Free Will, Love, And Anger. *Ideas y Valores*, 58(141), 169–189.
- Pereboom, D. (2014). *Personal Relationships and Meaning in Life*. Oxford University Press.
- Pickard, H. (2013). Irrational blame. *Analysis*, 73(4), 613–626. <https://doi.org/10.1093/analys/ant075>
- Pickard, Hanna. (2011). Responsibility Without Blame: Empathy and the Effective Treatment of Personality Disorder. *Philosophy, Psychiatry, and Psychology*, 18(3), 209–224.
- Pickard, Hanna. (2013). Responsibility without blame: Philosophical reflections on clinical practice. *Oxford Handbook of Philosophy of Psychiatry*, 1134–1154.
- Pickard, Hanna. (2014). Responsibility without Blame: Therapy, Philosophy, Law. *Prison Service Journal*, 213, 10.
- Pickard, Hanna. (2017a). Responsibility without Blame for Addiction. *Neuroethics*, 10(1), 169–180. <https://doi.org/10.1007/s12152-016-9295-2>
- Pickard, Hanna. (2017b). Responsibility without Blame for Addiction. *Neuroethics*, 10(1), 169–180. <https://doi.org/10.1007/s12152-016-9295-2>

- Ravenscroft, I. (2016). Folk Psychology as a Theory. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2016). Metaphysics Research Lab, Stanford University. Retrieved from <https://plato.stanford.edu/archives/fall2016/entries/folkpsych-theory/>
- Responsibility without Blame. (n.d.). Retrieved 30 November 2018, from <https://www.responsibilitywithoutblame.org/>
- Richman, K. A., & Bidshahri, R. (2017). Autism, theory of mind, and the reactive attitudes. *Bioethics*. <https://doi.org/10.1111/bioe.12370>
- Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology*, 21(2), 95.
- Rogers, C. R. (2013). A Theory of Therapy and Personality Change: As Developed in the Client-Centered Framework". *Perspectives in Abnormal Behavior: Pergamon General Psychology Series*, 341.
- Russell, P. (1992). Strawson's Way of Naturalizing Responsibility. *Ethics*, 102(2), 287–302.
- Russell, P. (2004). Responsibility and the Condition of Moral Sense. *Philosophical Topics*, 32(1–2), 287–305.
- Ryden, G., Ryden, E., & Hetta, J. (2008). Borderline personality disorder and Autism Spectrum Disorder in females - A cross-sectional study. *Clinical Neuropsychiatry*, 5(1), 22–30.
- Sagi, A., & Hoffman, M. L. (1976). Empathic distress in the newborn. *Developmental Psychology*, 12(2), 175–176.
- Scanlon, T. M. (2010). *Moral Dimensions: Permissibility, Meaning, Blame* (1 edition). Cambridge, Mass.: Harvard University Press.
- Schwitzgebel, E. (2017). On Not Distinguishing Too Finely Among One's Motivations. Retrieved 21 June 2017, from <http://schwitzsplinters.blogspot.com/2017/06/on-not-distinguishing-too-finely-among.html>
- Shabo, S. (2012). Incompatibilism and Personal Relationships: Another Look at Strawson's Objective Attitude. *Australasian Journal of Philosophy*, 90(1), 131–147.

- Sher, G. (2008). In Praise of Blame. *Philosophical Studies*, 137(1), 19–30.
- Shoemaker, D. (2007). Moral Address, Moral Responsibility, and the Boundaries of the Moral Community. *Ethics*, 118(1), 70–108. <https://doi.org/10.1086/521280>
- Shoemaker, D. (2011). Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility. *Ethics*, 121(3), 602–632. <https://doi.org/10.1086/659003>
- Shoemaker, D. (2015). *Responsibility from the Margins*. Oxford University Press.
- Shoemaker, D. (2017). Response-Dependent Responsibility; or, A Funny Thing Happened on the Way to Blame. *Philosophical Review*, 126(4), 481–527.
- Sie, Maureen. (2018). Sharing Responsibility: The Importance of Tokens of Appraisals to our Moral Practices. In K. Hutchison, C. Mackenzie, & M. Oshana (Eds.), *Social Dimensions of Moral Responsibility*. Oxford University Press.
- Singer, P. (1972). Famine, Affluence, and Morality. *Philosophy and Public Affairs*, 1(3), 229–243.
- Smetana, J. G. (1989). Toddlers' social interactions in the context of moral and conventional transgressions in the home. *Developmental Psychology*, 25(4), 499–508. <https://doi.org/10.1037/0012-1649.25.4.499>
- Smith, A. (2013). Moral Blame and Moral Protest. In D. J. Coates & N. A. Tognazzini (Eds.), *Blame: Its Nature and Norms*. Oxford University Press.
- Smith, A. M. (2007). On Being Responsible and Holding Responsible. *Journal of Ethics*, 11(4), 465–484.
- Snoek, A., Levy, N., & Kennett, J. (2016). Strong-willed but not successful: The importance of strategies in recovery from addiction. *Addictive Behaviors Reports*, 4, 102–107. <https://doi.org/10.1016/j.abrep.2016.09.002>
- Sripada, Chandra. (2015). Commentary on Kozuch and McKenna: Mental Illness, Moral Responsibility, and Expression of the Self. In D. D. Moseley & Gary, Gary (Eds.), *Philosophy and Psychiatry: Problems, Intersections and New Perspectives* (pp. 114–122). Routledge.

- Stout, N. (2016). Conversation, responsibility, and autism spectrum disorder. Retrieved 4 December 2017, from <https://philarchive.org>
- Strawson, P. F. (2008). *Freedom and resentment and other essays*. New York: Routledge.
- Strawson, Peter F. (1962). Freedom and Resentment. *Proceedings of the British Academy*, 48, 1–25.
- Talbert, M. (2008). Blame and responsiveness to reasons: are psychopath's blameworthy? *Pacific Philosophical Quarterly*, 89(4), 516–535.
- Todd, P. (2012). Manipulation and Moral Standing: An Argument for Incompatibilism. *Philosophers' Imprint*, 12(7).
- Todd, P. (2016). Strawson, Moral Responsibility, and the 'Order of Explanation': An Intervention. *Ethics*, 127(1).
- Vargas, M. (2013a). *Building Better Beings: A Theory of Moral Responsibility* (1 edition). Oxford: Oxford University Press.
- Vargas, M. (2013b). *Building Better Beings: A Theory of Moral Responsibility*. Oxford University Press.
- Vargas, M. (2013c). Situationism and Moral Responsibility: Free Will in Fragments. In T. Vierkant, J. Kiverstein, & A. Clark (Eds.), *Decomposing the Will*. Oxford Up.
- Walgrave, L. (2002). Not Punishing Children but Committing them to Restore. In I. Weijers & A. Duff (Eds.), *Punishing Juveniles: Principle and Critique*. Hart.
- Wallace, R. J. (1996). *Responsibility and the Moral Sentiments*. Harvard University Press.
- Watson. (2004). *Agency and Answerability*. Oxford University Press.
- Watson, G. (1993). Responsibility and the Limits of Evil: Variations on a Strawsonian Theme. In J. M. Fischer & M. Ravizza (Eds.), *Perspectives on Moral Responsibility* (pp. 119–148). Itaca: Cornell University Press.
- Weijers, I. (2002). A Pedagogical Perspective on Juvenile Justice. In I. Weijers & A. Duff (Eds.), *Punishing Juveniles: Principle and Critique*. Hart.
- Weijers, I., & Duff, A. (2002). *Punishing Juveniles: Principle and Critique*. Hart.

- White, S. W., Keonig, K., & Scahill, L. (2007). Social skills development in children with autism spectrum disorders: A review of the intervention research. *Journal of Autism and Developmental Disorders*, 37(10), 1858–1868.
- Williams, B. (1982). *Moral Luck*. Cambridge: Cambridge University Press.
- Xeniditis, K., Russell, A., & Murphy, D. (2001). Management of people with challenging behaviour. *Advances in Psychiatric Treatment*, 7(2), 109–116.
<https://doi.org/10.1192/apt.7.2.109>
- Zahavi, D., & Parnas, J. (2003). Conceptual problems in infantile autism research. *Journal of Consciousness Studies*.
- Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology*, 28(1), 126–136.
<https://doi.org/10.1037/0012-1649.28.1.126>